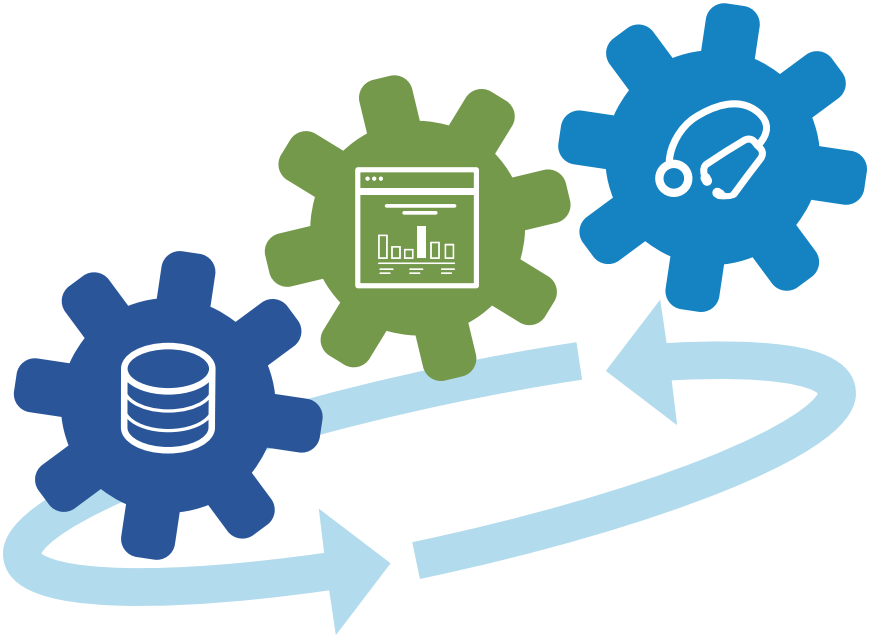
Handbook of eHealth Evaluation:

An Evidence-based Approach



EDITED BY

Francis Lau and Craig Kuziemsky

Handbook of eHealth Evaluation

Handbook of eHealth Evaluation: An Evidence-based Approach

EDITED BY
Francis Lau and Craig Kuziemsky

University of Victoria

Victoria, British Columbia
Canada
2016

Published by University of Victoria
Victoria, British Columbia Canada V8P 5C2
press@uvic.ca

Printed in Canada

Book design by Rayola Creative

Library and Archives Canada Cataloguing in Publication

Handbook of eHealth evaluation : an evidence-based approach / edited by Francis Lau and Craig Kuziemsky.

Includes bibliographical references and index.
Issued in print and electronic formats.
ISBN 978-1-55058-601-5 (paperback).--ISBN 978-1-55058-602-2 (pdf).--ISBN 978-1-55058-603-9 (epub)

1. Medical care--Information technology--Evaluation--Handbooks, manuals, etc. I. Lau, Francis Yin Yee, 1955-, editor II. Kuziemsky, Craig, 1970-, editor III. Title: eHealth evaluation.

R858.H35 2016 610.285'57 C2016-906658-4
C2016-906659-2

Contents

List of Tables and Figures

Tables

Figures

Preface

Why this Handbook?

The overall aim of this handbook is to provide a practical guide on the evaluation of eHealth systems. Over the years, we have seen a steady growth in the number and type of eHealth systems being adopted in different healthcare settings. Proponents of these systems claim eHealth can improve the quality of care provided, leading to better provider performance and health outcomes. Yet the evidence for such claims is mixed thus far, with some studies demonstrating benefits, others showing little to no impact, and some settings being even worse off than before. Understandably, there are now increasing pressures on government agencies and health organizations to demonstrate tangible return on value for the significant eHealth investments made.

Despite the growing importance and need to evaluate eHealth systems, there are relatively few formal courses available from post-secondary educational institutions on how to plan, conduct, report and appraise eHealth evaluation studies. Most educational institutions that offer degree programs related to health research, administration and services would typically include eHealth evaluation as part of their health research methods or program evaluation courses. Of those that offer health informatics degree programs, only some have eHealth evaluation as a full self-contained course. For institutions that offer eHealth evaluation as either a course or a topic within a course, the choice of textbooks and reference materials can vary greatly depending on what is available and the preference of the instructors.

To date, there have been just a few books published on eHealth evaluation. Notable examples are the reference texts edited by van Gennip and Talmon (1994), Anderson and Aydin (2005), and Friedman and Wyatt (2006), as well as the handbook written by Brender (2006). Aside from these, we are not aware of other major reference texts published in the last 10 years focused solely on this topic. Yet during this period we have witnessed an exponential growth in the number of published journal articles and government reports on eHealth evaluation. These publications often contain descriptions of different evaluation approaches and/or field studies on the design, implementation, use and effects of particular eHealth systems in specific settings. Overall, what seems lacking is a reference text that brings together these diverse approaches, studies and lessons as a coherent body of literature on the current state of knowledge in eHealth evaluation in a manner that is both rigorous and practical.

With the increasing use of eHealth systems and the growing demand to demonstrate their value, there is a strong case to be made to incorporate eHealth evaluation as part of the adoption process in order to generate the empirical evidence needed. Given the lack of current reference texts on eHealth

evaluation, we believe it is both necessary and timely to publish an up-to-date resource that can help those involved with eHealth evaluation in healthcare settings. Rather than publishing an academic textbook in the traditional manner, we have opted for a handbook in the form of a freely available electronic book or e-book. Compared to a conventional text, we believe such a freely available e-book can better serve as a more flexible, updatable and practical guide for those who need to plan, conduct, report and appraise eHealth evaluation in the field setting.

Who is it for?

This handbook is intended as a primary resource or a supplementary resource to textbooks on eHealth for students enrolled in courses related to eHealth evaluation. This handbook is also intended for individuals who are involved with the planning, design, implementation, use, support and assessment of eHealth systems in different healthcare settings. These individuals may be managers, analysts, developers, providers and trainees who are involved with some aspects of eHealth systems as part of their day-to-day work. In large organizations some of these individuals may have dedicated roles in eHealth evaluation. But often we expect them to be responsible for aspects of eHealth planning, design, implementation and support, with evaluation assigned as an afterthought or an adjunct role on the side.

The varied audience identified above suggests that this e-book is written for individuals who are not experts in eHealth evaluation but are expected to engage in such assessment activities in their own workplaces. In fact, much of the content in this handbook can be considered introductory in nature. This is to ensure those who are relatively new to the subject can gain a basic understanding of the current state of eHealth evaluation approaches, studies and findings, and can see how this knowledge could be applied and interpreted within their own settings.

At the same time, this handbook can also be a useful resource for individuals who are already familiar with eHealth evaluation. In particular, the handbook provides a systematic overview of the different evaluation approaches with case examples that have been applied and reported for a wide range of eHealth systems across different healthcare settings. As such, the handbook can serve as a reference text on details regarding particular evaluation approaches and the current state of knowledge in selected eHealth domains covered as case examples.

Francis Lau and Craig Kuziemsky
Editors

References

Anderson, J. G., & Aydin, C. E. (Eds.). (2005). *Evaluating the organizational impact of healthcare information systems* (2nd ed.). New York: Springer Science+Business Media, Inc.

Brender, J. (2006). *Handbook of evaluation methods for health informatics.* San Diego, CA: Elsevier Academic Press, U.S.

Friedman, C. P., & Wyatt, J. C. (2006). *Evaluation methods in medical informatics* (2nd ed.). New York: Springer Verlag, Inc.

van Gennip, E. M. S. J., & Talmon, J. L. (Eds.). (1995). *Assessment and evaluation of information technologies in medicine.* Amsterdam: IOS Press.

Acknowledgements

This handbook could not have come to fruition without the dedicated efforts of all the co-authors who have so graciously contributed their time and energy to produce the content seen here. The concept of writing a freely available e-book on eHealth evaluation is not without risk as it is unclear how the e-book will be received by the eHealth community. Yet all of the co-authors have enthusiastically embraced the idea of electronic self-publishing for which we are deeply indebted. Their names are already listed in the contributors section at the end of this handbook and thus will not be repeated here. We wish to thank Morgan Price who created the graphic design for the e-book cover. The support of the University of Victoria library staff under the direction of Inba Kehoe must also be acknowledged, as their diligent work was crucial in compiling the final e-book version and making it available on the university's website. Lastly a special thanks to James Allen, the relentless copy editor, whose meticulous attention to detail has supported the creation of an e-book of such fine quality.

Introduction

Francis Lau and Craig Kuziemsky

What is eHealth?

eHealth is an overarching term that refers to the use of information and communication technology (ICT) in the healthcare sector. Despite being a widely used and popular term there is no single universally agreed-upon definition of eHealth. At the dawn of the 21st century, an editorial on eHealth published in an online journal broadly defined the term as follows:

> eHealth is an emerging field in the interaction of medical informatics, public health and business, referring to health services and information delivered or enhanced through the Internet and related technologies. In a broad sense, the term characterizes not only a technical development, but also a state-of-mind, a way of thinking, an attitude, and a commitment for networked, global thinking, to improve health care locally, regionally, and worldwide by using information and communication technology. (Eysenbach, 2001, p. e20)

According to a scoping review by Pagliari et al. (2005) on the definition and meaning of eHealth, the term first appeared in year 2000 and has since become widely used. Of the 387 relevant articles these authors reviewed in 154 different journals, the most common usages were related to information technology (IT) and telemedicine, with an emphasis on the communicative aspects through networks and the Internet. The definitions they found varied widely in terms of the functions, stakeholders, contexts and theoretical issues involved.

In a systematic review on eHealth studies by Oh, Rizo, Enkin, and Jadad (2005), 51 definitions were found in 430 journals and 1,158 websites. All of the definitions mentioned health and technology. Most included varying aspects of stakeholders, their attitudes, the role of place and distance, and the expected benefits from eHealth. For health it usually referred to care processes rather than outcomes. For technology it was seen as both an enabling tool for a healthcare process or service and also as the resource itself such as a health information website. Moreover, there was an overwhelming sense of optimism in the definitions.

It is important to note that even now the term eHealth is used differently across countries. Here are examples of how the term eHealth is being used in Canada, the United States, Europe and Australia:

- Canada: eHealth is defined by Health Canada as the application of ICT in the healthcare sector with the electronic health record (EHR) as the basic building block (Health Canada, n.d.). Canadian jurisdictions have all used eHealth to refer to a broad range of ICT-based systems, services and resources in their business and IT plans. These include the electronic medical record (EMR), the personal health record (PHR), consumer health, telehealth/telemedicine, and public health surveillance. Note that in the Canadian context EHR includes information from laboratory and drug information systems, diagnostic imaging repositories, provider and patient registries, telehealth applications, and public health surveillance made available through privacy-protected interoperable platforms (Infoway, n.d.). Other terms that have also been used are electronic health information systems (Infoway, 2004) and more recently digital health (Infoway, 2016).

- United States: Both the terms health IT and eHealth are in common use. For instance, the Office of the National Coordinator for Health Information Technology (ONC) HealthIT.gov website section for patients and families explains that health IT "refers to technologies and tools that allow health care professionals and patients to store, share, and analyze health information" (ONC, n.d.). Examples of health IT listed include EHR and PHR that are used to store and share one's electronic health information. The ONC website also has a section on consumer eHealth programs which are intended to support ONC efforts to empower individuals to improve their health and healthcare through the use of health IT. Examples of eHealth programs include the Meaningful Use Incentives, Blue Button, Sharecare and Innovation Challenges (ONC, 2015).

- Europe: The European Commission (2012) defines eHealth as "the use of ICT in health products, services and processes combined with organisational change in healthcare systems and new skills, in order to improve health of citizens, efficiency and productivity in healthcare delivery, and the economic and social value of health" (p. 3, footnote 1). Examples are ICT-supported "interaction between patients and health-service providers, institution-to-institution transmission of data, or peer-to-peer communication between patients and/or health professionals" to assist in disease prevention, diagnosis, treatment and follow-up (p. 3, footnote 1). Of particular interest is the inclusion of wearable and portable personal health systems collectively referred to as mHealth.

- Australia: The National E-Health Transition Authority (NEHTA) defines eHealth as "electronically connecting up the points of care so that health information can be shared securely" (NEHTA, n.d.). One example is the My Health Record System, with such products as the shared health

summary, discharge summary, specialist letter, eReferral, and prescription and dispense records that are accessible through the Web-based national consumer portal.

We should point out that, while some regard eHealth as being the same as health informatics, we believe the two are fundamentally different concepts. As described earlier, eHealth is broadly defined as the use of ICT-based systems, services and resources as an enabler in managing health. In contrast, we view health informatics as an academic discipline that deals with the science and practice of health information with respect to its meaning, capture, organization, retrieval, communication and use in decision-making. Since much of the health information is electronic in nature, health informatics also deals with the underlying ICT systems that support the health information in use.

What is eHealth Evaluation?

The *Merriam-Webster Dictionary* (n.d.) defines evaluation as an act to "judge the value or condition of (something) in a careful and thoughtful way." By extension, we can define eHealth evaluation as an act to assess whether an eHealth system is functioning and producing the effects as expected. In this context, the eHealth system can be any ICT-based application, service or resource used by organizations, providers, patients or consumers in managing health. Here the concept of health refers to one's physical and mental condition, and its management refers to a wide range of health services and information resources used to maintain or improve one's state of well-being. Note that an eHealth system covers not only the technical ICT artefact but also the socio-organizational and environmental factors and processes that influence its behaviours.

The scope of eHealth evaluation can cover the entire life cycle, which spans the planning, design, implementation, use, and maintenance of the eHealth system over time. Depending on the life cycle stage being evaluated there can be different questions raised. For instance, in the planning stage of an eHealth system, one may evaluate whether the intended system is aligned with the organization's overall strategy, or if an adequate governance process is in place for the sharing of sensitive patient information. In the design stage one may evaluate whether the specifications of the system have been met in terms of its features and behaviour. In the implementation stage one may evaluate whether the deployment of the system is on time and within budget. In the use stage one may evaluate the extent to which the system is used and its impact on provider performance, health outcomes and economic return. In the maintenance stage one may evaluate how well the system is being supported and adapted to accommodate the changing needs of the organization over time.

Different eHealth evaluation approaches have been described in the literature ranging from randomized controlled trials (RCTs), qualitative studies, to usability engineering. These approaches all have unique philosophical and

methodological assumptions, leading to confusion as to when and how a particular approach should be applied and the implications involved. Some also regard eHealth evaluation as a form of research that is only relevant to those in academia. Our position is that eHealth evaluation should be scientifically rigorous, relevant to practice, and feasible to conduct in routine settings. By rigorous it means the approach should be credible and defensible. By relevant it means the problem being addressed should be important to the stakeholders. By feasible it means the design should be practical and achievable within a reasonable time frame using reasonable resources.

In their evaluation textbook, Friedman and Wyatt (2006, pp. 25–27) introduced the notion of an evaluation mindset with the following characteristics to distinguish it from research:

- Tailor the study to the problem, ensuring questions that are relevant to stakeholders are being addressed.
- Collect data useful for making decisions, focusing on data from processes that are relevant to decision-makers.
- Look for intended and unintended effects, assuming the effects of an eHealth system cannot be known in advance.
- Study the system while it is under development and after it is deployed, thus acknowledging the dynamic nature of an eHealth system where its effects can change over time.
- Study the system in the laboratory and in the field, thereby assessing the performance and effects of an eHealth system in both simulated and natural settings.
- Go beyond the developer's point of view, ensuring the perspectives of different stakeholders who are affected by the eHealth system are taken into account.
- Take the environment into account, understanding the surroundings in which the eHealth system resides.
- Let the key issues emerge over time, understanding the need for time passage before some issues become evident.
- Be methodologically broad and eclectic, recognizing the need for and importance of different approaches when planning, conducting and appraising an evaluation study.

In other words, eHealth evaluation should be considered in all endeavours related to an eHealth system because of the significant time and resources required to adopt and adapt these systems. Therefore it is important to find out whether and how much such effort has led to tangible improvement in one's performance and/or outcomes. In addition, there is an opportunity cost associated with investing in eHealth systems since that investment could be spent elsewhere, for example to reduce surgical wait times by increasing the volume of surgeries performed. Within the current climate of fiscal restraint in the health systems of many jurisdictions, there has to be a strong business case to justify the deployment of eHealth investments.

Thus far, eHealth evaluation studies are often conducted and reported by academic and leading health institutions that have made significant investments in eHealth systems and expert resources to improve their provider performance and health outcomes. While in recent years we have seen increased interest from health organizations in general to engage in eHealth evaluation, what appears to be missing are the necessary eHealth infrastructures and expertise to tackle such activities. By infrastructures we mean the ability to capture and extract the types of clinical and operational data needed to perform the evaluation. By expertise we mean the know-how of the different approaches used in evaluation. Therefore, some form of guidance is needed for stakeholders to engage in eHealth evaluation in a rigorous, relevant and pragmatic fashion. We offer this handbook as one source of such guidance.

What is in this Handbook?

This handbook presents the science and practice of eHealth evaluation based on empirical evidence gathered over many years within the health informatics discipline. The handbook describes different approaches used to evaluate the planning, design, implementation, use and impact of eHealth systems in different health settings. It also provides a snapshot of the current state of knowledge on the consequences of opting for eHealth systems with respect to their effects and implications on provider performance and health outcomes.

The science part of this handbook covers the conceptual foundations of and methodological details in eHealth evaluation. Conceptual foundations refer to the theories, models and frameworks that have been used as organizing schemes and mental roadmaps by eHealth practitioners to illuminate and clarify the makeup, behaviour and effects of eHealth systems beyond that of a technical artefact. Methodological details refer to the different approaches and methodologies that have been used to evaluate eHealth systems. Collectively they provide a rich set of tried and proven methods that can be readily applied or adapted for use by eHealth practitioners responsible for the evaluation of specific eHealth systems.

The practice part covers the ground-level application of the scientific eHealth evaluation approaches described in Parts I and II of the handbook, through the

presentation of a set of published case examples in Part III. These case studies provide a summary of the current state of evidence in selected eHealth systems and domains, and how the evaluation studies were designed, conducted and reported. Part IV of the handbook covers the future of eHealth evaluation. It describes the need to build intellectual capacity as a way of advancing the field by ensuring eHealth practitioners are well versed in the science and practice of eHealth evaluation. Also of importance is the need for a more strategic view of eHealth evaluation within the larger healthcare system to be successful.

This handbook has been written as an open electronic reference text or e-book that is to be freely available to students and practitioners wishing to learn about eHealth evaluation or apply the content in their workplace. This e-book is a "living book" in that the co-authors can add such content as new reviews, evaluation methods and case studies as they become available over time. An online learning community is also being considered depending on whether there is sufficient interest from the co-authors and the eHealth communities.

Note that throughout this handbook there are numerous terms mentioned in the form of acronyms and abbreviations. Rather than repeating the full spellings of these terms every time they are mentioned in the chapters, we have opted for the short form and provided a glossary of the acronyms and abbreviations at the end of the handbook (pp. 473–477).

References

European Commission. (2012, December 6). *eHealth action plan 2012-2010: Innovative healthcare for the 21st century.* Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions. Brussels: Author. Retrieved from http://ec.europa.eu/health/ehealth/docs/com_2012_736_en.pdf

Eysenbach, G. (2001). What is e-health? *Journal of Medical Internet Research*, *3*(2), e20.

Health Canada. (n.d.). *Health care system – eHealth*. Ottawa: Author. Retrieved from http://www.hc-sc.gc.ca/hcs-sss/ehealth-esante/index-eng.php

Infoway. (n.d.). *Electronic health records – Electronic health record components*. Toronto: Author. Retrieved from https://www.infoway-inforoute.ca/en/solutions/electronic-health-records

Infoway. (2004). *Annual report, 2003-2004*. Corporate plan summary, 2004-2005. Toronto: Author. Retrieved from https://www.infoway-inforoute.ca/en/component/edocman/153-annual-report-2003-2004-corporate-plan-summary-2004-2005/view-document?Itemid=101

Infoway. (2016). *Summary corporate plan, 2016-2017*. Toronto: Author. Retrieved from https://www.infoway-inforoute.ca/en/component/edocman/2858-summary-corporate-plan-2016-2017/view-document?Itemid=101

Merriam-Webster. (n.d.). Evaluation. *Merriam-Webster Dictionary*. Retrieved from http://www.merriam-webster.com/

National E-Health Transition Authority (NEHTA). (n.d.). *What is eHealth?* Sydney, Australia: Author. Retrieved from http://www.nehta.gov.au/get-started-with-ehealth/what-is-ehealth.

Office of the National Coordinator for Health Information Technology (ONC). (n.d.). *Health IT for you.* Washington, DC: Department of Health and Human Services. Retrieved from https://www.healthit.gov/sites/default/files/consumerfactsheetfinal.pdf

Office of the National Coordinator for Health Information Technology (ONC). (2015). *Consumer eHealth program.* Washington, DC: Department of Health and Human Services. Retrieved from https://www.healthit.gov/policy-researchers-implementers/consumer-ehealth-program

Oh, H., Rizo, C., Enkin, M., & Jadad, A. (2005). What is eHealth (3): A systematic review of published definitions. *Journal of Medical Internet Research, 7*(1), e1.

Pagliari, C., Sloan, D., Gregor, P., Sullivan, F., Detmer, D., Kahan, J. P., Ortwijn, W., & MacGillivray, S. (2005). What is eHealth (4): A scoping exercise to map the field. *Journal of Medical Internet Research, 7*(1), e9.

Part I

Conceptual Foundations

Chapter 1

Need for Evidence, Frameworks and Guidance

Francis Lau

1.1 Introduction

Over the years, a variety of countries and subnational jurisdictions have made significant investments in eHealth systems with the expectation that their adoption can lead to dramatic improvements in provider performance and health outcomes. With this increasing movement toward eHealth systems there is a consequent need for empirical evidence to demonstrate there are tangible benefits produced from these systems. Such evidence is imporant to establish the return on investment and value, as well as to guide future eHealth investment and adoption decisions.

Thus far the evidence on tangible eHealth benefits has been mixed. In light of these conflicting results, conceptual frameworks are needed as organizing schemes to help make sense of the evidence on eHealth benefits. In particular, it is important to appreciate the underlying assumptions and motivations governing an evaluation and its findings so that future eHealth investment and adoption decisions can be better informed. Along with the need for conceptual frameworks to make sense of the growing eHealth evidence base, there is also an increasing demand to provide best practice guidance in eHealth evaluation approaches to ensure there is both rigour and relevance in the planning, conduct, reporting and appraisal of eHealth evaluation studies.

This chapter describes the challenges associated with eHealth evaluation, and the need for empirical evidence, conceptual frameworks and practice guidance to help us make sense of eHealth evaluation. Six different frameworks that constitute the remaining chapters in Part I of this handbook are then outlined.

1.2 Evaluation Challenges

There are three types of challenges to be considered when navigating the eHealth evaluation landscape. These are the definition of eHealth itself, one's perspective of eHealth systems, and the approaches used to study eHealth systems. These challenges are elaborated below.

1.2.1 The Challenge of Definition

The field of eHealth is replete with jargons, acronyms and conflicting descriptions that can be incomprehensible to the uninitiated. For instance, eHealth is defined by some countries as the application of Information and Communication Technology (ICT) in health. It is a term often seen in the Canadian and European literature. On the other hand, Health Information Technology (HIT) is also a term used to describe the use of ICT in health especially in the United States. The terms EHR (Electronic Health Record) and EMR (Electronic Medical Record) can have different meanings depending on the countries in which they are used. In the United States, EHR and EMR are used interchangeably to mean electronic records that store patient data in health organizations. However, in Canada EMR refers specifically to electronic patient records in a physician's office.

The term EHR can also be ambiguous as to what it contains. According to the Institute of Medicine, an EHR has four core functions: health information, data storage, order entry (i.e., computerized provider/physician order entry, or CPOE), results management, and decision support (Blumenthal et al., 2006). Sometimes it may also include patient support, electronic communication and reporting, and population health management. Even CPOE can be ambiguous as it may or may not include decision support functions. The challenge with eHealth definitions, then, is that there are often implicit, multiple and conflicting meanings. Thus, when reviewing the evidence on eHealth design, adoption and impacts, one needs to understand what eHealth system or function is involved, how it is defined, and where and how it is used.

1.2.2 The Challenge of Perspective

The type of eHealth system and/or function being evaluated, the health setting involved, and the evaluation focus are important considerations that influence how various stakeholders perceive a system with respect to its purpose, role and value. Knowing the eHealth system and/or function involved – such as a CPOE with clinical decision support (CDS) – is important as it identifies what is being evaluated. Knowing the health setting is important since it embodies the type of care and services, as well as organizational practices, that influence how a system is adopted. Knowing the focus is to reduce medication errors with CDS is important as it identifies the value proposition being evaluated. Often the challenge with eHealth perspective is that the descriptions of the system, setting and focus are incomplete in the evaluation design and reporting. This lack of detail makes it difficult to determine the significance of the study findings and their relevance to one's own situation. For example, in studies of CPOE with CDS

in the form of automated alerts, it is often unclear how the alerts are generated, to whom they are directed, and whether a response is required. For a setting such as a primary care practice it is often unclear whether the site is a hospital outpatient department, a community-based clinic or a group practice. Some studies focus on such multiple benefit measures as provider productivity, care coordination and patient safety, which render it difficult to decide whether the system has led to an overall benefit. It is often left up to the consumer of evaluation study findings to tease out such detail to determine the importance, relevance and applicability of the evidence reported.

1.2.3 The Challenge of Approach

A plethora of scientific, psychosocial and business approaches have been used to evaluate eHealth systems. Often the philosophical stance of the evaluator influences the approach chosen. On one end of the spectrum there are experimental methods such as the randomized controlled trial (RCT) used to compare two or more groups for quantifiable changes from an eHealth system as the intervention. At the other end are descriptive methods such as case studies used to explore and understand the interactions between an eHealth system and its users. The choice of benefit measures selected, the type of data collected and the analytical techniques used can all affect the study results. In contrast to controlled studies that strive for statistical and clinical significance in the outcome measures, descriptive studies offer explanations of the observed changes as they unfold in the naturalistic setting. In addition, there are economic evaluation methods that examine the relationships between the costs and return of an investment, and simulation methods that model changes based on a set of input parameters and analytical algorithms.

The challenge, then, is that one needs to know the principles behind the different approaches in order to plan, execute, and appraise eHealth evaluation studies. Often the quality of these studies varies depending on the rigour of the design and the method applied. Moreover, the use of different outcome measures can make it difficult to aggregate findings across studies. Finally, the timing of studies in relation to implementation and use will influence impacts which may or may not be realized during the study period due to time lag effects.

1.3 Making Sense of eHealth Evaluation

The growing number of eHealth systems being deployed engenders a growing need for new empirical evidence to demonstrate the value of these systems and to guide future eHealth investment and adoption decisions. Conceptual frameworks are needed to help make sense of the evidence produced from eHealth evaluation studies. Practice guidance is needed to ensure these studies are scientifically rigorous and relevant to practice.

1.3.1 The Need for Evidence

The current state of evidence on eHealth benefits is diverse, complex, mixed and even contradictory at times. The evidence is diverse since eHealth evaluation studies are done on a variety of topics with different perspectives, contexts, purposes, questions, systems, settings, methods and measures. It is complex as the studies often have different foci and vary in their methodological rigour, which can lead to results that are difficult to interpret and generalize to other settings. The evidence is often mixed in that the same type of system can have either similar or different results across studies. There can be multiple results within a study that are simultaneously positive, neutral and negative. Even the reviews that aggregate individual studies can be contradictory for a given type of system in terms of its overall impacts and benefits.

To illustrate, a number of Canadian eHealth evaluation studies have reported notable benefits from the adoption of EMR systems (O'Reilly, Holbrook, Blackhouse, Troyan, & Goeree, 2012) and drug information systems (Fernandes et al., 2011; Deloitte, 2010). Yet in their 2009-2010 performance audit reports, the Auditor General of Canada and six provincial auditors offices raised questions on whether there was sufficient value for money on Canadian EHR investments (Office of the Auditor General of Canada [OAG], 2010). Similar mixed findings appear in other countries. In the United Kingdom, progress toward an EHR for every patient has fallen short of expectations, and the scope of the National Programme for IT has been reduced significantly in recent years but without any reduction in cost (National Audit Office [NAO], 2011). In the United States, early 21st century savings from health IT were projected to be $81 billion annually (Hillestead et al., 2005). Yet overall results in the U.S. have been mixed thus far. Kellerman and Jones (2013) surmised the causes to be a combination of sluggish health IT adoption, poor interoperability and usability, and an inability of organizations to re-engineer their care processes to reap the available benefits. Others have argued the factors that lead to tangible eHealth benefits are highly complex, context-specific and not easily transferable among organizations (Payne et al., 2013).

Despite the mixed findings observed to date, there is some evidence to suggest that under the right conditions, the adoption of eHealth systems are correlated with clinical and health system benefits, with notable improvements in care process, health outcomes and economic return (Lau, Price, & Bassi, 2015). Presently this evidence is stronger in care process improvement than in health outcomes, and the positive economic return is only based on a small set of published studies. Given the current societal trend toward an even greater degree of eHealth adoption and innovation in the foreseeable future, the question is no longer whether eHealth can demonstrate benefits, but under what circumstances can eHealth benefits be realized and how should implementation efforts be applied to address factors and processes that maximize such benefits.

1.3.2 The Need for Frameworks

In light of the evaluation challenges described earlier, some type of organizing scheme is needed to help make sense of eHealth systems and evaluation findings. Over the years, different conceptual frameworks have been described in the health informatics and information systems literature. For example, Kaplan (2001) advocated the use of such social and behavioural theories as social interactionism to understand the complex interplay of ICT within specific social and organizational contexts. Orlikowski and Iacono (2001) described the nominal, computational, tool, proxy and ensemble views as different conceptualizations of the ICT artefact in the minds of those involved with information systems.

In their review of evaluation frameworks for health information systems, Yusof, Papazafeiropoulou, Paul, and Stergioulas (2008) identified a number of evaluation challenges, examples of evaluation themes, and three types of frameworks that have been reported in eHealth literature. For evaluation challenges, one has to take into account the why, who, when, what and how questions upon undertaking an evaluation study:

- Why refers to the purpose of the evaluation.
- Who refers to the stakeholders and perspectives being represented.
- When refers to the stage in the system adoption life cycle.
- What refers to the type of system and/or function being evaluated.
- How refers to the evaluation methods used.

For evaluation themes, examples of topics covered include reviews of the impact of clinical decision support systems (CDSS) on physician performance and patient outcomes, the importance of human factors in eHealth system design and implementation, and human and socio-organizational aspects of eHealth adoption. The three types of evaluation frameworks reported were those based on generic factors, system development life cycle, and sociotechnical systems. Examples of generic factors are those related to the eHealth system, its users and the social-functional environment. Examples of system development life cycle are the stages of exploration, validity, functionality and impact. Examples of sociotechnical systems are the work practices of such related network elements as people, organizational processes, tools, machines and documents.

It can be seen that the types of conceptual frameworks reported in the eHealth literature vary considerably in terms of their underlying assumptions, purpose and scope, conceptual dimensions, and the level and choice of measures used. In this context, underlying assumptions are the philosophical stance of the evaluator and his or her worldview (i.e., subjective versus objective). Purpose and scope are the intent of the framework and the health domain that it covers. Conceptual di-

mensions are the components and relationships that make up the framework. Level and choice of measures are the attributes that are used to describe and quantify the framework dimensions. Later in this chapter, six examples of conceptual frameworks from the eHealth literature are introduced that have been used to describe, understand and explain the technical, human and organizational dimensions of eHealth systems and their sociotechnical consequences. These frameworks are then described in detail in Part I of this handbook.

1.3.3 The Need for Guidance

The term "evidence-based health informatics" first appeared in 1990s as part of the evidence-based medicine movement. Since that time, different groups have worked to advance the field by incorporating the principle of evidence-based practice into their health informatics teaching and learning. Notable efforts included the working groups of the University for Health Sciences, Medical Informatics and Technology (UMIT), International Medical Informatics Association (IMIA), and European Federation of Medical Informatics (EFMI), with their collective output called the Declaration of Innsbruck that laid the foundation of evidence-based health informatics and eHealth evaluation as a recognized and growing area of study (Rigby et al., 2013).

While much progress has been made thus far, Ammenwerth (2015) detailed a number of challenges that still remain. These include the quality of evaluation studies, publication biases, the reporting quality of evaluation studies, the identification of published evaluation studies, the need for systematic reviews and meta-analyses, training in eHealth evaluation, the translation of evidence into practice and post-market surveillance. From the challenges identified by this author, it is clear that eHealth evaluation practice guidance is needed in multiple areas and at multiple levels. First, guidance on multiple evaluation approaches is needed to examine the planning, design, adoption and impact of the myriad of eHealth systems that are available. Second, guidance is needed to ensure the quality of the evaluation study findings and reporting. Third, guidance is needed to educate and train individuals and organizations in the science and practice of eHealth evaluation.

In this regard, the methodological actions of the UMIT-IMIA-EFMI working groups that followed their Declaration of Innsbruck have been particularly fruitful in moving the field of eHealth evaluation forward (Rigby et al., 2013). These actions include the introduction of guidelines for good eHealth evaluation practice, standards for reporting of eHealth evaluation studies, an inventory of eHealth evaluation studies, good eHealth evaluation curricula and training, systematic reviews and meta-analyses of eHealth evaluation studies, usability guidelines for eHealth applications, and performance indicators for eHealth interventions. In aggregation, all of these outputs are intended to increase the rigour and relevance of eHealth evaluation practice, promote the generation and reporting of empirical evidence on the value of eHealth systems, and in-

crease the intellectual capacity in eHealth evaluation as a legitimate field of study. In Part II of this handbook, different approaches from the eHealth literature that have been applied to design, conduct, report and appraise eHealth evaluation studies are described.

1.4 The Conceptual Foundations

In Part I of this handbook, the chapters that follow describe six empirical frameworks that have been used to make sense of eHealth systems and their evaluation. These frameworks serve a similar purpose in that they provide an organizing scheme or mental roadmap for eHealth practitioners to conceptualize, describe and predict the factors and processes that influence the design, implementation, use and effect of eHealth systems in a given health setting. At the same time, these frameworks are different from each other in terms of their scope, the factors and processes involved, and their intended usage. The six frameworks covered in chapters 2 through 7 are introduced below.

- Benefits Evaluation (BE) Framework (Lau, Hagens, & Muttitt, 2007) – This framework describes the success of eHealth system adoption as being dependent on three conceptual dimensions: the quality of the information, technology and support; the degree of its usage and user satisfaction; and the net benefits in terms of care quality, access and productivity. Note that in this framework, organizational and contextual factors are considered out of scope.

- Clinical Adoption (CA) Framework (Lau, Price, & Keshavjee, 2011) – This framework extends the BE Framework to include organizational and contextual factors that influence the overall success of eHealth system adoption in a health setting. This framework has three conceptual dimensions made up of micro-, meso- and macro-level factors, respectively. The micro-level factors are the elements described in the BE Framework. The meso-level factors refer to elements related to people, organization and implementation. The macro-level factors refer broadly to elements related to policy, standards, funding and trends in the environment.

- Clinical Adoption Meta-Model (CAMM) (Price & Lau, 2014) – This framework provides a dynamic process view of eHealth system adoption over time. The framework is made up of four conceptual dimensions of availability, use, behaviour and outcomes. The basic premise is that for successful adoption to occur the eHealth system must first be made available to those who need it. Once available, the system has to be used by the intended users as part of their day-to-day work. The ongoing use of the system should gradually

lead to observable behavioural change in how users do their work. Over time, the behavioural change brought on by ongoing use of the system by users should produce the intended change in health outcomes.

- eHealth Economic Evaluation Framework (Bassi & Lau, 2013) – This framework provides an organizing scheme for the key elements to be considered when planning, conducting, reporting and appraising eHealth economic evaluation studies. These framework elements cover perspective, options, time frame, costs, outcomes and analysis of options. Each element is made up of a number of choices that need to be selected and defined when describing the study.

- Pragmatic HIT Evaluation Framework (Warren, Pollock, White, & Day, 2011) – This framework builds on the BE Framework and a few others to explain the factors and processes that influence the overall success of eHealth system adoption. The framework is multidimensional and adaptive in nature. The multidimensional aspect ensures the inclusion of multiple viewpoints and measures, especially from those who are impacted by the system. The adaptive aspect allows an iterative design where one can reflect on and adjust the evaluation design and measures as data are being collected and analyzed over time. The framework includes a set of domains called criteria pool made up of a number of distinct factors and processes for considerations when planning an evaluation study. These criteria are work and communication patterns, organizational culture, safety and quality, clinical effectiveness, IT system integrity, usability, vendor factors, project management, participant experience and leadership, and governance.

- Holistic eHealth Value Framework (Lau, Price, & Bassi, 2015) – This framework builds on the BE, CA and CAMM Frameworks by incorporating their key elements into a higher-level conceptual framework for defining eHealth system success. The framework is made up of the conceptual dimensions of investment, adoption, value and lag time, which interact with each other dynamically over time to produce specific eHealth impacts and benefits. The investment dimension has factors related to direct and indirect investments. The adoption dimension has micro-, meso- and macro-level factors described in the BE and CA Frameworks. The value dimension is conceptualized as a two-dimensional table with productivity, access and care quality in three rows and care process, health outcomes and economic return in three columns. The lag time dimension has adoption lag time and impact lag time, which

take into account the time needed for the eHealth system to be implemented, used and to produce the intended effects.

1.5 Summary

This chapter explained the challenges in eHealth evaluation and the need for empirical evidence, conceptual frameworks and practice guidance to make sense of the field. The six frameworks used in eHealth evaluation that are the topics in the remaining chapters of Part I of this handbook were then introduced.

References

Ammenwerth, E. (2015). Evidence-based health informatics: How do we know what we know? *Methods of Information in Medicine, 54*(4), 298–307.

Bassi, J., & Lau, F. (2013). Measuring value for money: A scoping review on economic evaluation of health information systems. *Journal of American Medical Informatics Association, 20*(4), 792–801.

Blumenthal, D., DesRoches, C., Donelan, K., Ferris, T., Jha, A., Kaushal, R., … Shield, A. (2006). *Health information technology in the United States: the information base for progress.* Princeton, NJ: Robert Wood Johnson Foundation.

Deloitte. (2010). *National impacts of generation 2 drug information systems.* Technical Report, September 2010. Toronto: Canada Health Infoway. Retrieved from https://www.infoway-inforoute.ca/index.php/en/component/edocman/resources/reports/331-national-impact-of-generation-2-drug-information-systems-technical-report

Fernandes, O. A., Lee, A. W., Wong, G., Harrison, J., Wong, M., & Colquhoun, M. (2011). What is the impact of a centralized provincial drug profile viewer on the quality and efficiency of patient admission medication reconciliation? A randomized controlled trial. *Canadian Journal of Hospital Pharmacy, 64*(1), 85.

Hillestad, R., Bigelow, J., Bower, A., Girosi, F., Meili, R., Scoville, R., & Taylor, R. (2005). Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. *Health Affairs, 24*(5), 1103–1117.

Kaplan, B. (2001). Evaluating informatics applications — some alternative approaches: theory, social interactionism, and call for methodological pluralism. *International Journal of Medical Informatics, 64*(1), 39–58.

Kellerman, A. L., & Jones, S. S. (2013). What it will take to achieve the as-yet-unfulfilled promises of health information technology. *Health Affairs, 32*(1), 63–68.

Lau, F., Hagens, S., & Muttitt, S. (2007). A proposed benefits evaluation framework for health information systems in Canada. *Healthcare Quarterly, 10*(1), 112–118.

Lau, F., Price, M., & Keshavjee, K. (2011). From benefits evaluation to clinical adoption: Making sense of health information system success in Canada. *Healthcare Quarterly, 14*(1), 39–45.

Lau, F., Price, M., & Bassi, J. (2015). Toward a coordinated electronic health record (EHR) strategy for Canada. In A. S. Carson, J. Dixon, & K. R. Nossal (Eds.), *Toward a healthcare strategy for Canadians* (pp. 111–134). Kingston, ON: McGill-Queens University Press.

National Audit Office. (2011). *The national programme for IT in the NHS: an update on the delivery of detailed care records systems*. London: Author. Retrieved from https://www.nao.org.uk/report/the-national-programme-for-it-in-the-nhs-an-update-on-the-delivery-of-detailed-care-records-systems/

Office of the Auditor General of Canada [OAG]. (2010, April). *Electronic health records in Canada – An overview of federal and provincial audit reports*. Ottawa: Author. Retrieved from http://www.oag-bvg.gc.ca/internet/docs/parl_oag_201004_07_e.pdf

O'Reilly, D., Holbrook, A., Blackhouse, G., Troyan, S., & Goeree, R. (2012). Cost-effectiveness of a shared computerized decision support system for diabetes linked to electronic medical records. *Journal of the American Medical Informatics Association, 19*(3), 341–345.

Orlikowski, W. J., & Iacono, C. S. (2001). Research commentary: Desperately seeking the "IT" in IT research – A call to theorizing the IT artefact. *Information Systems Research, 12*(2), 121–134.

Payne, T. H., Bates, D. W., Berner, E. S., Bernstam, E. V., Covvey, H. D., Frisse, M. E., … Ozbolt, J. (2013). Healthcare information technology and economics. *Journal of the American Medical Informatics Association*, *20*(2), 212–217.

Price, M., & Lau, F. (2014). The clinical adoption meta-model: a temporal meta-model describing the clinical adoption of health information systems. *BMC Medical Informatics and Decision Making, 14*, 43. Retrieved from http://www.biomedical.com/1472-6947/14/43

Rigby, M., Ammenwerth, E., Beuscart-Zephir, M.- C., Brender, J., Hypponen, H., Melia, S., Nykänen, P., Talmon, J., & de Keizer, N. (2013). Evidence-based health informatics: 10 years of efforts to promote the principle. *IMIA Yearbook of Medical Informatics, 2013*, 34–46.

Warren, J., Pollock, M., White, S., & Day, K. (2011). *Health IT evaluation framework*. Wellington, NZ: Ministry of Health.

Yusof, M. M., Papazafeiropoulou, A., Paul, R. J., & Stergioulas, L. K. (2008). Investigating evaluation frameworks for health information systems. *International Journal of Medical Informatics, 77*(6), 377–385.

Chapter 2

Benefits Evaluation Framework

Francis Lau, Simon Hagens, Jennifer Zelmer

2.1 Introduction

The Benefits Evaluation (BE) Framework was published in 2006 as the result of a collective effort between Canada Health Infoway (Infoway) and a group of health informaticians. Infoway is an independent not-for-profit corporation with the mission to accelerate the development, adoption and effective use of digital health innovations in Canada. The health informaticians were a group of researchers and practitioners known for their work in health information technology (HIT) and health systems data analysis. These individuals were engaged by Infoway to be members of an expert advisory panel providing input to the pan-Canadian benefits evaluation program being established by Infoway at the time. The expert advisory panel consisted of David Bates, Francis Lau, Nikki Shaw, Robyn Tamblyn, Richard Scott, Michael Wolfson, Anne McFarlane and Doreen Neville.

At the time in Canada, the increased focus on evaluation of eHealth, both nationally and in the provinces and territories, reflected similar interest internationally. There was an increasing demand for evidence-informed investments, for information to drive optimization, and for accountability at project completion (Hagens, Zelmer, Frazer, Gheorghiu, & Leaver, 2015). The expert advisory panel recognized that a framework was a necessary step to convert that interest into focused action and results.

The intent of the BE Framework was to provide a high-level conceptual scheme to guide eHealth evaluation efforts to be undertaken by the respective jurisdictions and investment programs in Canada. An initial draft of the BE Framework was produced by Francis Lau, Simon Hagens, and Sarah Muttitt in early 2005. It was then reviewed by the expert panel members for feedback. A revised version of the framework was produced in fall of 2005, and published

in *Healthcare Quarterly* in 2007 (Lau, Hagens, & Muttitt, 2007). Supporting the BE Framework, the expert panel also led the development of a set of indicator guides for specific technologies and some complementary tools to allow broad application of the framework. Since its publication, the BE Framework has been applied and adapted by different jurisdictions, organizations and groups to guide eHealth evaluation initiatives across Canada and elsewhere.

This chapter describes the conceptual foundations of the BE Framework and the six dimensions that made up the framework. We then review the use of this framework over the years and its implications on eHealth evaluation for healthcare organizations.

2.2 Conceptual Foundations

The BE Framework is based on earlier work by DeLone and McLean (1992, 2003) in measuring the success of information systems (IS) in different settings, the systematic review by van der Meijden, Tange, Troost, and Hasman (2003) on the determinants of success in inpatient clinical information systems (CIS), and the synthesis of evaluation findings from published systematic reviews in health information systems (HIS) by Lau (2006) and Lau, Kuziemsky, Price, and Gardner (2010). These published works are summarized below.

2.2.1 Information Systems Success Model

The original IS Success Model published by DeLone and McLean in 1992 was derived from an analysis of 180 conceptual and empirical IS studies in different field and laboratory settings. The original model has six dimensions of IS success defined as system quality, information quality, use, user satisfaction, individual impact, and organizational impact (Figure 2.1). Each of these dimensions represents a distinct construct of "success" that can be examined by a number of quantitative or qualitative measures. Examples of these measures for the six IS success dimensions are listed as follows:

- *System quality* – ease of use; convenience of access; system accuracy and flexibility; response time
- *Information quality* – accuracy; reliability; relevance; usefulness; understandability; readability
- *Use* – amount/duration of use; number of inquiries; connection time; number of records accessed
- *User satisfaction* – overall satisfaction; enjoyment; software and decision-making satisfaction

- *Individual impact* – accurate interpretation; decision effectiveness, confidence and quality

- *Organizational impact* – staff and cost reductions; productivity gains; increased revenues and sales

In 2003, DeLone and McLean updated the IS Success Model based on empirical findings from another 285 journal papers and conference proceedings published between 1992 and 2002 that validated, examined or cited the original model. In the updated model a service quality dimension was added, and the individual and organizational impact dimensions were combined as a single construct called net benefits (Figure 2.2). The addition of service quality reflected the need for organizations to recognize the provision of IS service support beyond the technology as a determinant of IS success. Examples of *service quality measures* are *staff reliability, empathy and responsiveness*. On the other hand, the net benefits dimension was chosen to simplify the otherwise increasing number and type of impacts being reported such as group, industry and societal impacts. Also the inclusion of the word "net" in net benefits was intentional, as it emphasized the overall need to achieve positive impacts that outweigh any disadvantages in order for the IS to be considered successful.

The IS Success Model by DeLone and McLean is one of the most widely cited conceptual models that describe the success of IS as a multidimensional construct. It is also one of the few models that have been empirically validated in numerous independent laboratory and field evaluation studies across different educational, business and healthcare settings.

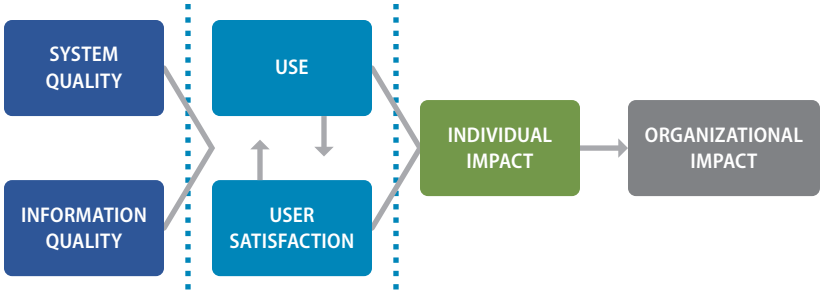


Figure 2.1. IS success model.

Note. From "Information systems success: The quest for the dependent variable," by W. H. DeLone and E. R. McLean, 1992, *Information Systems Research, 3*(1), p. 87. Copyright 1992 by INFORMS, http://www.informs.org. Reprinted with permission.

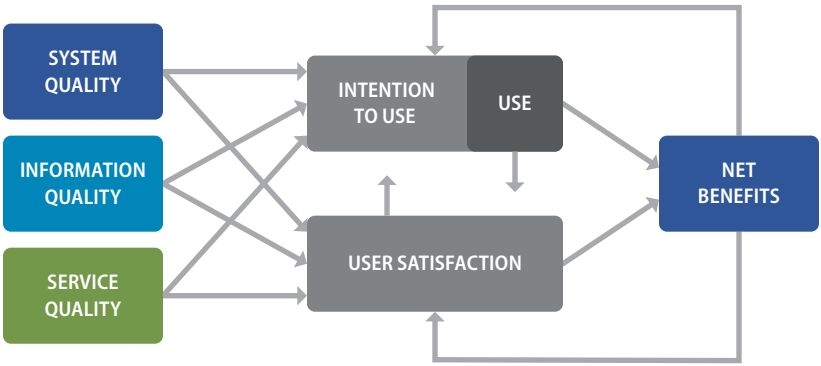


Figure 2.2. Updated IS success model.

Note. From "The DeLone and McLean model of information systems success: A ten-year update," by W. H. DeLone and E. R. McLean, 2003, *Journal of Management Information Systems, 19*(4), p. 24. Copyright 2003 by Taylor & Francis. Reprinted with permission.

2.2.2 Clinical Information Systems Success Model

Van der Meijden et al. (2003) conducted a literature review on evaluation studies published from 1991 to 2001 that identified attributes used to examine the success of inpatient clinical information systems (CIS). The review used the IS Success Model developed by DeLone and McLean as the framework to determine whether it could correctly categorize the reported attributes from the evaluation studies. In total, 33 studies describing 29 different CIS were included in the review, and 50 attributes identified from these studies were mapped to the six IS success dimensions (Table 2.1). In addition, 16 attributes related to system development, implementation, and organizational aspects were identified as contingent factors outside of the six dimensions in the IS Success Model (Table 2.2).

Table 2.1

Attributes of CIS Success Factors

System Quality Attributes	Information Quality Attributes	Usage Attributes	User Satisfaction Attributes	Individual Impact Attributes	Organizational Impact Attributes
• Ease of use – (record keeping time), • Response time, • Time savings, • Intrinsic features creating extra work, • Perceived ease of use, • Usability, • Availability, • Ease of learning, • Rigidity of system – (built-in rules), • Reliability, • Security, • Easy access to help, • Data accuracy	• Completeness, • Accuracy of data, • Legibility, • Timeliness, • Perceived usefulness, • Availability, • Comprehensive, • Consistency, • Reliability, • Format	• Number of entries, • Frequency of use, • Duration of use, • Self-reported usage, • Location of data entry, • Frequency of use of specific functions	• User satisfaction, • Attitude, • User friendliness, • Expectations, • Competence in computers	• Changed clinical work patterns, • Direct benefits, • Changed documentation habits – (more administrative tasks, time of day for documenting, documentation frequency), • Information use – (information recall, accurate interpretation, integration of information / overview, information awareness), • Efficiency and effectiveness of work, • Job satisfaction	• Communication and collaboration, • Impact on patient care, • Costs – (time savings, reduction of staff, number of procedures reduced)

Note. From "Determinants of success of clinical information systems: A literature review," by M. J. van der Meijden, H. J. Tange, J. Troost, and A. Hasman, 2003, *Journal of the American Medical Informatics Association, 10*(3), p. 239. Copyright 2003 by Oxford University Press, on behalf of the American Medical Informatics Association. Adapted with permission.

Table 2.2

Attributes of Contingent Factors

System Development Attributes	Implementation Attributes	Organizational Aspects Attributes
• User involvement • Redesign work practices • Reconstruction of content / format • Technical limitations	• Communication (frequency, two way) • Training • Priorities chosen • Technical support • User involvement	• Organizational culture – (control and decision-making, management support, professional values, collaboration / communication) • Support and maintenance • Champions • Rewards

Note. From "Determinants of success of clinical information systems: A literature review," by M.J. van der Meijden, H. J. Tange, J. Troost, and A. Hasman, 2003, *Journal of the American Medical Informatics Association, 10*(3), p. 241. Copyright by Oxford University Press, on behalf of the American Medical Informatics Association. Adapted with permission.

Since its publication in 2003, the CIS Success Model by van der Meijden and colleagues (2003) has been widely cited and applied in eHealth evaluation studies. The CIS Success Model can be considered an extension of the original IS Success Model in that it recognizes the influence and importance of contingent factors related to the system development, implementation and organizational aspects that were not included in the original model.

2.2.3 Synthesis of Health Information System Reviews

Lau (2006) examined 28 systematic reviews of health information system (HIS) evaluation studies published between 1996 and 2005. From an initial synthesis on 21 of the published reviews pertaining to clinical information systems/tools and telehealth/telemedicine evaluation studies, Lau identified 60 empirical evaluation measures in 20 distinct categories of success factors based on the six IS success dimensions in the revised DeLone and MacLean model (i.e., system, information and service quality, use and user satisfaction, and net benefits). These empirical evaluation measures were reconciled with the success measures reported in the original and revised DeLone and MacLean models, as well as the attributes identified in the van der Meijden et al. model. Additional findings from the Lau review that were supplementary to the BE Framework included the clinical domains, study designs and evaluation measures used in the evaluation studies. These findings provided an initial empirical evidence base for the potential application of the BE Framework dimensions, categories and measures (Lau, 2006). Selected findings for 14 of the initial 21 systematic reviews examined are shown in Table 2.3. See also the separate additional references section for Table 2.3.

Table 2.3
Summary of 14 Systematic Review Articles on HIS Field Evaluation Studies

Authors	Topic	Design	Evaluation Metrics
Ammenwerth and de Keizer (2004)	Health info systems, evaluation	1,035 studies	Journal, type, location, method, focus
Balas et al. (1998)	Clinical info systems	98 RCT	Process and outcome of care
Balas et al. (1996)	Diabetes management	15 CT	48 outcome measures reported
Cramer et al. (2003)	Computerized health evidence delivery	57 RCT, 10 SR	Process of care, patient health, others
Delpierre et al. (2004)	Patient record systems	26 studies	Practice, quality of care, satisfaction
Garg et al. (2005)	CDSS	100 CT	Performance and outcome
Kaushal et al. (2003)	CPOE, CDSS medication safety	12 trials	Behaviours, med errors, adverse events
Kawamoto et al. (2005)	CDSS	70 RCT	Improved clinical practice
Mitchell and Sullivan (2001)	CDSS in primary care	89 CT, B/A	Performance and outcomes
Montgomery and Fahey (1998)	Hypertension management	7 RCT	Performance, improved blood pressure
Sullivan and Mitchell (1995)	Computerized primary care consultation	30 studies	Consult time, preventions, satisfaction
van der Loos et al. (1995)	Health information systems in diffusion	108 studies	Structure, process, outcome measures
van der Meijden et al. (2003)	Inpatient clinical info systems	33 studies	Quality, use and impact
Walton et al. (1999)	Optimum drug dosage	18 trials	Effect size, relative % difference

Legend: CDSS – clinical decision support system; RCT – randomized control trial; CT – controlled trial; SR – systematic review; B/A – before/after; TS – time series; EMR – electronic medical record; DS – decision support

Note. From "Increasing the rigor of health information system studies through systematic reviews," by F. Lau, 2006, a presentation to *11th International Symposium on Health Information Management Research (iSHIMR)*, Halifax, Nova Scotia, Canada.

2.3 Benefits Evaluation Framework Dimensions

The BE Framework is based on all six dimensions of the revised DeLone and MacLean IS Success Model, which are system, information and service quality, use and user satisfaction, and net benefits. A total of 20 categories and 60 subcategories of evaluation measures are defined in the BE Framework. They are based on the measures identified in the van der Meijden et al. (2003) CIS Success Model and the Lau et al. (2010) HIS review synthesis. In the BE Framework, the net benefits are further grouped into three subcategories of care quality, access and productivity. These subcategories are from the original benefits measure-

ment framework defined by Infoway to determine the impact of digital health broadly on national healthcare renewal priorities (Infoway, 2005).

When creating the BE Framework, Infoway recognized the importance of organizational and contextual factors on the adoption and impact of eHealth systems. However, these factors were considered out-of-scope at the time in order to reduce the complexity of the framework. The scope was also tailored to increase its acceptance by stakeholder organizations, as many of the eHealth project teams who would be overseeing evaluation were not well positioned to investigate and report on the broader issues. The BE Framework is shown in Figure 2.3. Note that there are other measures in the IS and CIS success models that are not in the BE Framework. They were excluded for such pragmatic reasons as the perceived subjective nature of the data and the difficulty in their collection.

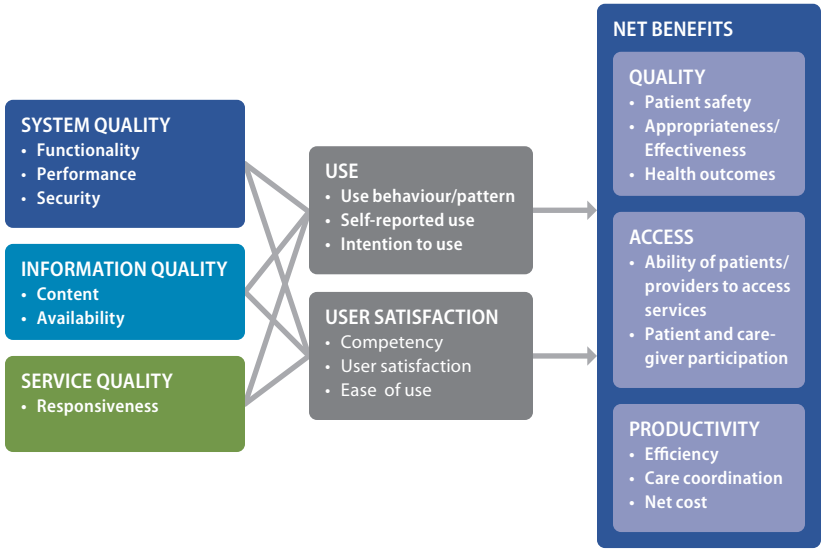


Figure 2.3. Infoway benefits evaluation (BE) framework.

Note. Copyright 2016 by Canada Health Infoway Inc., http://www.infoway-inforoute.ca. Reprinted with permission.

2.3.1 Health Information Technology Quality

There are three HIT quality dimensions, namely system, information, and service.

> *System quality* refers to the technical aspects of the HIT and has three categories of measures on system functionality, performance and security. Functionality covers the type and level of HIT features present such as order entry with decision support for reminders and alerts. Performance covers the technical behaviour of the HIT in terms of its accessibility, reliability and response time. Security

covers the ability to protect the integrity and use of the data captured, and to ensure only authorized access to the HIT.

Information quality refers to the characteristics of the data in the system and has two categories on the quality of the content and its availability. Content covers the accuracy, reliability, completeness and comprehension of the data. Availability covers the timeliness of accessing the data when and where needed.

Service quality refers to HIT implementation, training and ongoing support by staff and has one category on responsiveness. Examples of responsiveness are the extent and adequacy of user training and technical support available. Not included are service empathy and assurance from the IS success model which were considered too subjective to evaluate at that time. Note that for each of the BE Framework dimensions and categories there are further breakdowns into subcategories and measures. See section 2.3.4 for a complete list of the defined HIT quality measures.

2.3.2 Use and User Satisfaction

The use dimension in the BE Framework has three categories which are usage behaviour and pattern, self-reported use, and intention to use. Usage behaviour and pattern cover actual HIT usage in terms of type, frequency, duration, location and flexibility. One example is the volume of medication orders entered by providers on the nursing units in a given time period. Self-reported use covers perceived HIT usage reported by users in terms of type, frequency, duration, location and flexibility. Intention to use is the proportion of and factors causing non-users of an implemented HIT to become active users of the system. The satisfaction dimension has three categories, namely competency, user satisfaction, and ease of use. Competency covers the knowledge, skills and experience of the users in the HIT. User satisfaction covers the extent to which the users feel gratified from using the HIT to accomplish their tasks. Ease of use covers the extent to which the users feel the HIT is both easy to learn and easy to use.

2.3.3 Net Benefits

The net benefits dimension has three categories of measures on care quality, access and productivity, respectively. Care quality has three subcategories: patient safety, appropriateness and effectiveness, and health outcomes. Patient safety includes adverse events, prevention, surveillance, and risk management. Appropriateness includes the adherence and compliance to benchmarks, policy or practice standards, and self-reported practices or practice profiles captured in the system. Effectiveness includes continuity of care with individuals or local/dispersed teams and referral of services. Health outcomes include short-term

clinical outcomes and longer-term change in the health status of patients attributable to HIT interventions.

Access has two subcategories that cover the ability of the patient to access care, which includes enabling access to care through technology (e.g., videoconferencing) and driving improvements in access (e.g., wait time information systems), and the extent of patient/caregiver participation in these services. Productivity has three subcategories: efficiency, care coordination, and net cost. Efficiency includes resource use, output and care continuity improvement, and health systems management capability. Care coordination includes care provision by teams and continuity of care across settings. Net cost includes monetary avoidance, reduction and saving.

2.3.4 Summary of Benefit Evaluation Measures

The BE Framework dimensions, categories, subcategories and measures are summarized in Table 2.4. Note that these are suggested measures only, and are not an exhaustive list of measures reported in the literature. Healthcare organizations may choose to adopt these measures or adapt and extend the list to include new measures to suit their needs.

Table 2.4
Summary of BE Measures

Dimension	Category	Subcategories and Definitions of Measures
System	Functionality	Type and level of features available (e.g., order entry and decision support)
	Performance	Accessibility (remote and availability), reliability (up/down time) and system response time
	Security	Type and level of features available
Information	Content	Accuracy, relevance, completeness and comprehension
	Availability	Timeliness, reliability and consistency of data when and where needed
Service	Responsiveness	Extent and adequacy of implementation, training and ongoing support available
Use	User behaviour and pattern	Type, frequency, duration, location and flexibility of actual usage
	Self-reported use	Type, frequency, duration, location and flexibility of perceived usage
	Intention to use	Reasons for current non-users to become users and proportion who do
Satisfaction	Competency	Knowledge, skills and experience of users in the HIS
	User satisfaction	Extent to which the users feel gratified from using the HIS
	Ease of use	User friendliness and learnability

Table 2.4
Summary of BE Measures

Dimension	Category	Subcategories and Definitions of Measures
Net benefits	Care Quality	Patient safety - preventable adverse events, near-misses and errors - surveillance in monitoring of specific populations for patterns and trends - reduction in patient risks and safety-related reportable adverse events
		• Patient safety a) preventable adverse events, near-misses and errors b) surveillance in monitoring of specific populations for patterns and trends c) reduction in patient risks and safety-related reportable adverse events
		• Appropriateness and effectiveness a) adherence and compliance with benchmark, policy or practice standards and guidelines b) self-reported practice or practice captured in the HIS c) immunization and testing and other relevant rates d) continuity of care, examples: • information, relational and management continuity • by individuals or multi-disciplinary or geographically dispersed teams • access to information and effectiveness of general practitioner and specialist referral
		• Health outcomes a) clinical outcomes b) change in health status attributable to eHealth interventions
	Access	• Ability of patients and providers to access services a) availability, diversity and consolidation of eHealth-enabled services b) timeliness, geographic, financial and cultural or linguistic c) removal of inequitable barriers (including affordability, acceptability and accommodation) • Patient and caregiver participation a) patients' self-management and access to their own information
	Productivity	• Efficiency a) provider resource use b) improvement short term outputs vs. inputs, and long term in care continuity c) improved health system management capability d) improved patient efficiency (e.g., more efficient scheduling of preoperative testing) e) non-monetary effects • Care coordination a) care provision by team b) continuity of care across continuum • Net cost a) monetary avoidance b) monetary reductions, actual/projected savings

Note. From "A proposed benefits evaluation framework for health information systems in Canada," by F. Lau, S. Hagens, and S. Muttitt, 2007, *Healthcare Quarterly, 10*(1), p. 115. Copyright 2007 by Longwoods™ Publishing Corp. Reprinted with permission.

2.4 Benefit Evaluation Framework Usage

Since its debut in 2006, the BE Framework has been applied, adapted and cited in different evaluation reports, reviews, studies and commentaries. In this section we describe the companion resources that were created along with the framework. Then we summarize evaluation studies conducted in Canada that applied, adapted and cited the framework, followed by studies from other countries. Last, we include an example of a survey tool that can be used to evaluate the adoption of eHealth systems.

2.4.1 Companion Resources

The BE Framework is helpful in describing factors that influence eHealth success. But there should also be guidance and resources in place to help practitioners apply the framework in specific field evaluation studies. Guidance can be in the form of suggested evaluation questions, methods, designs and measures that are appropriate for the type of eHealth system and adoption stage involved, as well as the logistics for collecting and analyzing the data needed in the study. Another form of guidance required relates to managing evaluation activities, from structuring stakeholder engagement and gaining buy-in, to finding skilled evaluators, overseeing studies, and communicating results. Resources can be in the form of sample evaluation study plans, data collection tools, best practices in eHealth evaluation, completed evaluation reports and published peer-reviewed evaluation studies. As part of the initial release of the BE Framework in 2006, Infoway commissioned leading experts to develop indicator guides and compiled a BE Indicators Technical Report (Infoway, 2006) and a System and Use Assessment (SUA) survey tool (Infoway, 2006) as two companion resources. These resources were developed in collaboration with the Infoway BE expert advisory panel, eight subject matter experts, and two consultant teams.

The 2006 BE Indicators Technical Report (Infoway, 2006) includes a detailed description of the BE Framework, suggested evaluation questions, indicators and measures for specific eHealth programs, criteria for selecting appropriate BE indicators, and examples of tools and methods used in completed evaluation studies. The report covers six program areas, which are diagnostic imaging, drug information systems, laboratory information systems, public health systems, interoperable Electronic Health Records (iEHRs) and telehealth. These were some of the core initial investment programs funded by Infoway where it was necessary to assess tangible benefits to the jurisdictions and healthcare organizations as co-funders of these programs. Version 2.0 of the BE Indicators Technical Report was released in 2012 with expanded content (Infoway, 2012). The report still covers six program areas but laboratory information system has been merged with interoperable EHR as one section, and electronic medical records (EMR) for physician/nurse practitioner offices has been added as a new section. In Version 2.0 there are many more examples of published evaluation studies including those from Canadian jurisdictions and healthcare organizations. A BE planning template has also been added to facilitate the creation of a

practical evaluation plan for any eHealth system, and provide some of the practical guidance on managing evaluation activities. Since the publication of Version 2.0, additional program indicator sets and tools have been developed for telepathology, consumer health solutions and ambulatory EMR.

The SUA survey tool was introduced in 2006 as a multipart semi-structured questionnaire to collect information from users on the quality of the eHealth system and its usage in the organization. The questionnaire has since been adopted as a standardized Infoway survey tool to collect comparable information on the quality and use of eHealth systems being evaluated in Canada (Infoway, 2012). The SUA survey tool is aligned with the HIT quality, use and satisfaction dimensions of the BE Framework in terms of the questions used. The current version of this survey tool has eight sections of questions and guidance on how to administer the survey and analyze the results for reporting. These sections are on overall user satisfaction, system quality, information quality, service quality, public health surveillance, system usage, other comments, and demographic information. The survey can be adapted or expanded to include specific questions tailored to a particular eHealth system, such as the perceived accuracy of the images from the diagnostic imaging system being evaluated (Infoway, 2012).

2.4.2 Benefit Evaluation Framework Usage in Canada

Over the years, the BE Framework has been applied in over 50 evaluation studies across Canada. As examples, Table 2.5 shows 13 Canadian evaluation studies conducted over the past six years. See also the separate additional references section for Table 2.5. Six of these studies were related to telehealth, covering such clinical areas as ophthalmology, oncology and chronic disease management (British Columbia Ministry of Health [MOH], 2011a; B.C. MOH, 2011b; Gartner Inc., 2013; Praxia Information Intelligence & Gartner, Inc., 2010; Ernst & Young, 2014; Newfoundland and Labrador Centre for Health Information [NLCHI], 2010). Two studies covered drug information systems (Deloitte, 2010; Gartner Inc., 2013). Two studies covered diagnostic imaging systems (Gartner Inc., 2013; Hagens et al., 2009a). Two studies were on EMR systems for ambulatory and community care settings, respectively (PricewaterhouseCoopers [PWC], 2013; MOH, 2014). There was also one study each on vaccine inventory management (B.C. MOH, 2013), electronic occurrence reporting for patient safety (Elliot, 2014) and SNOMED (Systematized Nomenclature of Medicine) Clinical Terms (CT)[1] use in palliative care (Lau, 2010).

1 In 2014, the International Health Terminology Standards Development Organisation (IHTSDO) responsible for SNOMED CT officially changed the name so SNOMED CT no longer refers to Systematized Nomenclature of Medicine Clinical Terms, but rather just SNOMED Clinical Terms. It has become a trade name rather than an acronym.

Most of these evaluation studies focused on satisfaction, care quality, productivity and access dimensions of the BE Framework, with the addition of measures specific to eHealth systems as needed. Examples include turnaround time for imaging test results, patient travel time and cost, and SNOMED CT term coverage in palliative care. Most studies used mixed methods to collect and analyse data from multiple sources. Reported methods include survey, interview, literature review, service data analysis and modelling of benefit estimates. Reported data sources include provider and patient surveys, interview and focus group data, service utilization data, prior evaluation reports and published peer-reviewed evaluation studies and systematic reviews. Note that many of the evaluation studies were based on perceived benefits from providers and patients, or projected benefits based on model cost estimates.

The BE Framework has also been cited in a number of Canadian evaluation studies, commentaries and student reports. For instance, in their evaluation of a provincial drug information system, Mensink and Paterson (2010) adapted the use and satisfaction dimensions of the BE Framework to examine its adoption and evolution over time. Similarly Shachak et al. (2013) extended the HIT service quality dimension to include different end user support themes such as onsite technical and data quality support by knowledgeable staff. In their commentary on EHR success strategy, Nagle and Catford (2008) emphasized the need to incorporate benefits evaluation as a key component toward EHR success. O'Grady and colleagues (2009) discussed collaborative interactive adaptive technologies (e.g., social media). Six graduate-level theses that drew on the BE Framework have also been published. These include the evaluation studies on: a scanning digital prescriber order system by Alsharif (2012); end user support for EMR by Dow (2012); electronic occurrence reporting on patient safety by Elliot (2010); EMR implementation in an ambulatory clinic (Forland, 2008); a multidisciplinary cancer conferencing system detailed by Ghaznavi (2012); and characteristics of health information exchanges in literature (Ng, 2012).

Table 2.5

Examples of Canadian Evaluation Studies where the BE Framework was Applied

Authors	Setting	eHealth system	Evaluation Focus	Design/ Methods	Indicator/ Measures	Results
B.C. MOH (2011a)	Six health regions in British Columbia	Telehealth system for specialized oncology consults, provider education	Access to oncology service and provider education, travel time and cost	Patient and physician surveys, analysis of utilization	Consult/education service counts, travel time, patient and physician satisfaction	Interim results showed increased access, reduced travel time, high satisfaction level for patients and providers
B.C. MOH (2011b)	Two health regions in British Columbia	Telehealth system for ophthamology retinal screening	Telehealth quality, use, satisfaction, access, productivity, empowerment	Survey, pre/post service use	System function, info quality, usability, travel time, patient volume, satisfaction, change in # diabetic and retinal screening	100% satisfied with telehealth quality and use, some travel cost saving, fee code and improved scheduling to maximize service
B.C. MOH (2013)	Four health regions in British Columbia	Panorama vaccine inventory module	Productivity, module usability, adoption, support mechanisms	Survey and interview	Staff time efficiency, vaccine wastage cost and volume	Some benefits, below expectations, need to streamline steps, expand functions/use
Cousins and Baldwin (2014)	Ambulatory clinics in provincial region in British Columbia	eChart/EMR	Key performance indicators for eChart	Survey, chart review, focus group, document review	eChart quality, usage, satisfaction, patient flow, medication alerts, patient/ family experience	Overall satisfied – 42%, quality acceptable –system 50%, info 63%, productivity +/- 10%
Deloitte (2010)	Pan-Canadian	Generation 2 drug info systems	Expected benefits in quality and productivity, focus on safety	Prior evaluations, survey, interviews, utilization analysis, benefits modelling	Adverse drug events and admissions, med abuse, compliance, productivity	Estimated benefits $436m: quality $252m, productivity $184m

Table 2.5

Examples of Canadian Evaluation Studies where the BE Framework was Applied

Authors	Setting	eHealth system	Evaluation Focus	Design/ Methods	Indicator/ Measures	Results
Elliot et al. (2014)	One health region in Newfound-land and Labrador	Electronic occurrence reporting system (aka clinical safety reporting system)	Benefits and lessons	Mixed methods, pre/post design, surveys, interviews, focus groups, cases reported, project documents	Pre/post adoption cases reported, time to reporting, usability and satisfaction	Increased reporting, improved notification, satisfaction, issues in implementation
Ernst & Young (2014)	Pan-Canadian, based on 4 programs	Remote patient monitoring (RPM) systems	Expected benefits in care quality, access and productivity	Utilization data, literature review, interviews, surveys	Utilization, break-even, cost, caregiver burden, satisfaction, compliance	Moderate evidence on benefits especially larger scale programs, solutions emerging
Gartner, Inc. (2013)	British Columbia, province-wide	Diagnostic imaging, Drug info systems, telehealth systems	Estimated benefits in care quality, access and productivity	Estimates from pan-Canadian studies, B.C. data and interviews	Expected cost saving, productivity, patient transfer, satisfaction, adverse events, callbacks, medication abuse, compliance, travel time, access	Expected improvement in care quality, access and productivity in DI at $90m, DIS at $200m and telehealth at $15m
Hagens et al. (2009a, 2009b)	Pan-Canadian based on 4 provinces	Diagnostic imaging systems	Estimated benefits in quality, access and productivity	Mixed methods, pre/post adoption survey, utilization	Turnaround time, transfers, duplicate exams, productivity, communicat ion, cost per case	Estimated benefits in improved access 30-40%, efficiency $160-190m, turnaround time 41%, productivity 25-30% at $122-148m
Lau et al. (2010)	Palliative care program in one region	Palliative care info system (PCIS)	SNOMED CT quality and use in palliative care	Mixed methods, interviews, case analysis and system usability	SNOMED CT quality, use/satisfac-tion, care quality, productivity	Higher consistency with SNOMED encoded consult letter with better quality

Table 2.5
Examples of Canadian Evaluation Studies where the BE Framework was Applied

Authors	Setting	eHealth system	Evaluation Focus	Design/ Methods	Indicator/ Measures	Results
NLCHI (2010)	Province-wide, Newfound-land and Labrador	Telehealth systems for chronic disease manage-ment	Service access and patient empower-ment	Surveys, interviews, utilization, admin data analysis	Service utilization and access, travel time, cost, continuity, follow-up, satisfaction	Increased service and access, high satisfaction and improved service, capacity limit, privacy concerns
Praxia and Gartner, Inc. (2010)	Pan-Canadian	Telehealth systems	Benefits in care quality, access and productivity, use and satisfaction	Utilization analysis, survey, literature review, interviews, prior evaluation	Utilization, travel time, cost avoidance, satisfaction	Utilization, estimated cost avoidance $55m and travel $70m 2010, socio-technical issues
PwC (2013)	Pan-Canadian	Community based EMR systems	Estimated benefits in care quality, access and productivity	Literature review, interviews, benefit estimate modeling	Expected benefits in efficiency, safety, outcomes, utilization, interaction	Expected efficiency gain $177m, less adverse events and duplicate tests $123m

2.4.3 Benefit Evaluation Framework Usage in Other Countries

The BE Framework has also been adapted or cited by health informaticians from other countries in their eHealth evaluation work. In New Zealand, for example, Warren, Pollock, Day, Gu, and White (2011) and Warren, Gu, Day, and Pollock (2012) have incorporated the BE Framework as part of their standardized criteria pool of evaluation measures to be used selectively when evaluating eHealth systems. The criteria pool covers work and communication patterns, organizational culture, safety and quality, clinical effectiveness, IT system integrity, usability, vendor factors, project management, participant experience, and leadership and governance. Warren and colleagues advocated the use of action research to conduct evaluation based on a select set of evaluation measures from the criteria pool. This approach has been applied successfully in the evaluation of electronic referral systems (Gu, Warren, Day, Pollock, & White, 2012; Warren et al., 2012).

In their literature review of routine health information systems (RHIS) in low- and middle-income countries, Hotchkiss, Dianna, and Foreit (2012) examined nine conceptual frameworks including the BE Framework for adaptation to evaluate the performance of RHIS and their impact on health system functioning. Ahmadi, Rad, Nilashi, Ibrahim, and Almaee (2013) applied a fuzzy model called

Technique for Order Performance by Similarity to Ideal Solution (TOPSIS) to identify the 10 most important factors in hospital EMR adoption based on 23 factors derived from the BE Framework. In addition, the evaluation toolkit for health information exchange projects from the United States Agency for Healthcare Research and Quality references a number of the measures from the BE Indicators Technical Report (Infoway, 2006) as recommendations for U.S. health information exchange projects (Cusack, Hook, McGowan, Poon, & Atif, 2010). A summary on the use of the BE Framework by these authors is shown in Table 2.6. See also the separate additional references section for Table 2.6.

Table 2.6
eHealth Evaluation in Other Countries where the BE Framework was Mentioned

Authors	Setting	eHealth system	Evaluation Focus	Design/ Methods	Indicators/ Measures	Results
Ahmadi et al. (2013) Malaysia	Private hospital	EMR systems	Ranking of most important factors in BE Framework	Survey, modeling with fuzzy technique for order performance by similarity to ideal solution (TOPSIS)	Likert-scale surveys with 23 parameters in 6 dimensions	10 important factors were patient choice, use strategies, ease of use, use intent, safety, communication, template, downtime, cost savings/profits
Cusack et al. (2010) United States	Multiple provider groups and healthcare organizations	Health information exchange (HIE)	Evaluation toolkit used to create an evaluation plan for HIE projects	A step-by-step process to determine HIE project goals and feasible measures	Measures for the process of creating a HIE and types of data used; and clinical process and outcome measures for the value propostion of HIE	Example measures listed in Sections II and III that are drawn from the BE Technical Indicators Report (2006)
Gu et al. (2012) New Zealand	Two health regions	Electronic referral in colorectal domain	Comparing two knowledge engineering (KE) project approaches	Mixed methods comparison of two cases	Criteria pool based on BE Framework dimensions	BE Framework guided examination of development approach, KE products, uptake and acceptance

Table 2.6
eHealth Evaluation in Other Countries where the BE Framework was Mentioned

Authors	Setting	eHealth system	Evaluation Focus	Design/ Methods	Indicators/ Measures	Results
Hotchkiss et al. (2012) United States	Low/middle income countries	Routine health information systems (RHIS)	RHIS performance, evaluation issues, improving evidence base	Literature review on conceptual frameworks and RHIS studies on effectiveness	Conceptual frameworks linking RHIS investments with performance, as inputs, processes, outputs, outcomes	BE Framework was one of nine conceptual frameworks cited
Nguyen and Bakewell (2011) Australia	One service provider organization	HIS for aged care providers	Impact of HIS adoption for aged care providers	Case study approach with mixed methods	HIS quality, use, satisfaction, and net benefits	Cited BE Framework but used revised D&M IS success model
Warren et al. (2011) New Zealand	National health IT systems	National shared care planning for long term conditions	Creation of a health IT evaluation framework	Action research approach with mixed methods	Criteria pool of measures for selection in specific evaluation studies	BE Framework dimensions included as part of criteria pool
Warren et al. (2012) New Zealand	Four healthcare organizations	Electronic referral systems	Comparison of four system features, adoption and benefits	Mixed methods	16 domains selected from criteria pool of evaluation measures	BE Framework dimensions as part of criteria pool, reported as lessons learned

2.4.4 System and Use Assessment Survey Tool Usage

The System and Use Assessment (SUA) survey tool has been applied in different eHealth evaluation studies across Canada. Recent examples include the evaluation of teleophthalmology and vaccine inventory management systems (Ministry of Health [MOH], 2011, 2013) and eChart (Cousins & Baldwin, 2014) in British Columbia, shared EHR in a western jurisdiction (Kuhn & Lau, 2014), and the drug information system in Prince Edward Island (Prince Edward Island [P.E.I.], 2010). A summary of these evaluation studies and how the survey tool was applied is shown in Table 2.7. See also the separate additional references section for Table 2.7.

There are also evaluation studies where the SUA survey has been adapted or cited. For instance, one Canadian jurisdiction – Nova Scotia – adapted the SUA survey tool to include more specific questions in the evaluation of their interoperable EHR picture archival and communication (PAC) and diagnostic imaging (DI) systems (for details, see Newfoundland and Labrador Centre for Health Information [NLCHI], 2014). Many of these studies are also available on the

Canada Health Infoway website. Other Canadian researchers adapted the survey tool to examine the quality and use of physician office EMRs (Paterson et al., 2010). In the United States, Steis et al. (2012) adapted the survey tool to examine user satisfaction with an electronic dementia assessment tool. In Saudi Arabia, Bah et al. (2011) adapted the tool to determine the level and extent of EHR adoption in government hospitals.

Table 2.7
Canadian Evaluation Studies where the S&U Assessment Survey Tool was Applied

Authors	Setting	eHealth system	Evaluation Focus	Design/ Methods	Indicators/ Measures	Results
B.C. MOH (2011)	Two health regions	Telehealth system for ophthalmology retinal screening	Telehealth quality, use, satisfaction, access, productivity, empowerment	Survey, pre/post service use	System function, info quality, usability, travel time, patient volume, satisfaction, change in # diabetic and retinal screening	100% satisfied with telehealth quality and use, some travel cost saving, fee code and improved scheduling to maximize service
B.C. MOH (2013)	Four health regions	Panorama vaccine inventory module	Productivity, module usability, adoption, support mechanisms	Survey and interview	Staff time efficiency, vaccine wastage cost and volume	Some benefits, below expectations, need to streamline steps, expand functions/use
Cousins and Baldwin (2014)	Ambulatory clinics in provincial health authority	eChart/EMR	Key performance indicators for eChart	Survey, chart review, focus group, document review	eChart quality, usage, satisfaction, patient flow, medication alerts, patient/family experience	Overall satisfied – 42%, quality acceptable - system 50%, info 63%, productivity +/- 10%
Kuhn and Lau (2014)	A western jurisdiction	Web-based shared EHR system	Use, satisfaction and impact of EHR	Survey and system use log	Adoption level, user satisfaction, impact	Info sharing improved, usage increased, issues with access, workflow integration

Table 2.7
Canadian Evaluation Studies where the S&U Assessment Survey Tool was Applied

Authors	Setting	eHealth system	Evaluation Focus	Design/ Methods	Indicators/ Measures	Results
Eapen and Chapman (2015)	Southwest Ontario	Mobile interface to EHR viewer	usability, impact on quality of patient care and productivity of health care providers	Survey	Adoption, usability, perceived productivity and quality	Users perceived the mobile interface of Clinical-Connect as useful but were neutral about the ease of use
P.E.I. (2010)	Province-wide	Drug information system	Stakeholder benefits, patient outcomes	Survey, admin data review	System/info quality, satisfaction, use, efficiency, drug compliance/ use	Slow but increasing use and satisfaction, need more training/ support

2.5 Implications

The BE Framework has proved to be a helpful conceptual scheme in describing and understanding eHealth evaluation. The BE Indicators Report and the SUA survey tool have become useful resources for healthcare organizations to plan and conduct evaluation studies on specific eHealth systems. The published evaluation studies that incorporated the BE Framework have provided a growing empirical evidence base where such studies can be reported, compared and aggregated over time. That said, there are both conceptual and practical implications with the BE Framework that should be considered. These implications are described below.

2.5.1 Conceptual Implications

There are conceptual implications related to the BE Framework in terms of its scope, definition and perspective. For scope, the BE Framework has purposely excluded organizational and contextual factors to be manageable. Note that the IS success model by DeLone and McLean (1992, 2003) has also made no mention of organizational and contextual factors. There was an assumption in that work that the IS involved were mature and operational systems with a stable user base, which made adoption issues less central. Yet many healthcare organizations are continuing to adopt and/or adapt eHealth systems due to changing legislation, strategies and technologies. As such, organizational and contextual factors can have a great deal of influence on the success of these eHealth systems. This limitation is evident from the contingent factors identified in the CIS

review by van der Meijden et al. (2003) and in the published evaluation studies from Canada and elsewhere.

This gap was one of the drivers for the development of the complementary National Change Management (CM) Framework (Infoway, 2012). Infoway facilitated the development of this framework though the pan-Canadian Change Management Network, with the intent of providing projects with practical tools to successfully implement eHealth change. Measurement is at the centre of the framework, surrounded by governance and leadership, stakeholder engagement, communications, training and workflow analysis and integration. Infoway has encouraged the use of the BE and CM frameworks in concert.

For definition, while the BE Framework dimensions, categories and measures have been established from empirical evidence over time, they are still concepts that can be interpreted differently based on one's experience and understanding of the meaning of these terms. In addition, the evaluation measures in the BE Framework are not exhaustive in what can be measured when evaluating the adoption and impact of myriad eHealth systems in different healthcare settings. As such, the caveat is that the definition of concepts and measures can affect one's ability to capture key aspects of specific eHealth systems for reporting, comparison and aggregation as part of the growing eHealth evidence base.

For perspective, it should be made clear that benefits evaluation and eHealth success are concepts that are dependent on the views and intentions of the stakeholders involved. There are many questions concerning what is considered "success" including: Who defines success? Who benefits from success? What is the trade-off to achieve success? These are questions that need to be addressed early when planning the eHealth system and throughout its design, implementation and evaluation stages. In short, the BE Framework can be perceived differently according to the various perspectives of stakeholders.

2.5.2 Practical Implications

There are also practical implications with the BE Framework in terms of how it is applied in real-life settings. One question raised frequently is how one should apply the framework when planning an evaluation study in an organization. To do so, one needs to consider the intent of the evaluation with respect to its focus, feasibility and utility.

For focus, one should identify the most important questions to be addressed and prioritize them accordingly in the evaluation. The BE Framework has a rich set of measures covering different aspects of eHealth adoption and impact, but one should not attempt to include all of them within a single study. For instance, if the focus of a study is to demonstrate the ability of an eHealth system to reduce medication errors, then one should select only a few key patient safety measures such as the incidents of adverse drug events reported over two or more time periods for comparison.

For feasibility, one should determine the availability of the data for the measures needed in the evaluation, as well as the time, resources and expertise avail-

able to design the study, collect and analyze the data, and report on the findings. For example, randomized controlled trials are often considered the gold standard in evaluating healthcare interventions. Yet it may be infeasible for the organization that is implementing an eHealth system to conduct such a trial since it is still adjusting to the changes taking place with the system. Similarly, an organization may not have the baseline data needed or the expertise available to conduct evaluation studies. In these situations the organization has to decide how feasible it is to capture the data or acquire the expertise needed. Capacity to conduct evaluation is another feasibility consideration, as more complex evaluations may require specialized skill sets of evaluators, funding, leadership support or other inputs that are limiting factors for some organizations.

For utility, one needs to determine the extent to which the evaluation efforts and results can inform and influence change and be leveraged for added value. The planning and conduct of an evaluation study can be a major undertaking within an organization. Executive and staff commitment is necessary to ensure the results and issues arising from the study are addressed to reap the benefits to the system. To maximize the utility of an evaluation study and its findings, one should systematically document the effort and results in ways that allow its comparison with studies from other organizations, and aggregation as part of the evolving empirical evidence base.

2.6 Summary

This chapter described the BE Framework as a conceptual scheme for understanding eHealth results. The framework has six dimensions in system, information and service quality, use and satisfaction, and net benefits, but organizational and contextual factors are considered out-of-scope. Since its debut in 2006, the BE Framework has been applied, adapted and cited by different jurisdictions, organizations and groups in Canada and elsewhere as an overarching framework to plan, conduct and report eHealth evaluation studies. Additional studies continue to be published on a regular basis. Recognizing its limitations in addressing contexts, there is a growing evidence base in the use of the BE Framework to evaluate the success of eHealth systems across different healthcare settings.

References

Alsharif, S. (2012). *Evaluation of medication turnaround time following implementation of scanning digital prescriber order technology*. MSc Internship Report in Health Informatics, Dalhousie University, Halifax, NS.

Bah, S., Alharthi, H., Mahalli, A. A. E., Jabali, A., Al-Qahtani M., & Al-Kahtani, N. (2011). Annual survey on the level and extent of usage of electronic health records in government-related hospitals in eastern province, Saudi Arabia. *Perspectives in Health Information Management, 8*(Fall), 1b.

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research*, *3*(1), 60–95.

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*, *19*(4), 9–30.

Dow, R. (2012). *The impact of end-user support on electronic medical record success in Ontario primary care: A critical case study*. MSc thesis in Information, University of Toronto.

Elliott, P. G. (2010). *Evaluation of the implementation of an electronic occurrence reporting system at Eastern Health, Newfoundland and Labrador (Phase One)*. PhD dissertation in Medicine, Memorial University of Newfoundland, St. John's.

Forland, L. (2007). *Evaluating the implementation of an electronic medical record for a health organization-affiliated family practice clinic*. MSc thesis in Health Informatics. University of Victoria, Victoria, BC, Canada.

Ghaznavi, F. (2012). *Design and evaluation of a multidisciplinary cancer conferencing platform*. MSc thesis in Biomedical Engineering. Chalmers University of Technology and University of Toronto.

Hagens, S., Zelmer, J., Frazer, C., Gheorghiu, B., & Leaver, C. (2015, February). Valuing national effects of digital health investments: An applied method. Paper presented at *Driving quality in informatics: Fulfilling the promise*, an international conference addressing information technology and communication in health, Victoria, BC, Canada.

Infoway. (2005). *Canada Health Infoway Inc. corporate business plan 2005–06*. Toronto: Author.

Infoway. (2006). *Benefits evaluation survey process — System & use assessment survey*. Canada Health Infoway benefits evaluation indicators technical report. Version 1.0, September 2006. Toronto: Author. Retrieved from https://www.infoway-inforoute.ca/en/component/edocman/resources/toolkits/change-management/national-framework/monitoring-and-evaluation/resources-and-tools/991-benefits-evaluation-survey-process-system-use-assessment-survey

Infoway. (2012). *Benefits evaluation survey process — System & use assessment survey*. Canada Health Infoway benefits evaluation indicators technical report. Version 2.0, April 2012. Toronto: Author.

Kuhn, K., & Lau, F. (2014). Evaluation of a shared electronic health record. *Healthcare Quarterly*, *17*(1), 30–35.

Lau, F. (2006). Increasing the rigor of health information system studies through systematic reviews? In R. Abidi, P. Bath, & V. Keselj (Eds.), *Proceedings of 11th International Symposium on Health Information Management Research (ISHIMR)*. July 14 to 16, Halifax, NS, Canada.

Lau, F., Hagens, S., & Muttitt, S. (2007). A proposed benefits evaluation framework for health information systems in Canada. *Healthcare Quarterly, 10*(1), 112–118.

Lau, F., Kuziemsky, C., Price, M., & Gardner, J. (2010). A review on systematic reviews of health information system studies. *Journal of American Medical Informatics Association*, *17*(6), 637–645. doi: 10.1136/jamia.2010.004838

Mensink, N., & Paterson, G. (2010). The evolution and uptake of a drug information system: The case of a small Canadian province. In C. Safran, H. Marin, & S. Reti (Eds.), *MEDINFO 2010* (pp. 352–355). Amsterdam: IOS Press. doi: 10.3233/978-1-60750-588-4-352.

Nagle, L. M., & Catford, P. (2008). Toward a model of successful electronic health record adoption. *Healthcare Quarterly*, *11*(3), 84–91.

Newfoundland and Labrador Centre for Health Information [NLCHI]. (2014). Nova Scotia post-implementation survey and interview results, March 2014. Retrieved from https://www.infoway-inforoute.ca/en/component/edocman/resources/reports/benefits-evaluation/2099-atlantic-canada-iehr-labs-benefits-evaluation-project-phase-2-nova-scotia

Ng, Y. (2012). *Key characteristics of health information exchange: A scoping review*. MSc report in Health Informatics. University of Victoria, Victoria, BC, Canada.

O'Grady, L., Witterman, H., Bender, J. L., Urowitz, S., Wiljer, D., & Jadad, A. R. (2009). Measuring the impact of a moving target: Towards a dynamic framework for evaluating collaborative adaptive interactive technologies. *Journal of Medical Internet Research, 11*(2), e20.

Paterson, G., Shaw, N., Grant, A., Leonard, K., Delisle, E., Mitchell, S., ... Kraetschmer, N. (2010). A conceptual framework for analyzing how Canadian physicians are using electronic medical records in clinical care. In C. Safran, S. Reti, & H. F. Marin (Eds.), *Proceedings of the 13th World Congress on Medical Informatics* (pp. 141–145). Amsterdam: IOS Press.

Shachak, A., Montgomery, C., Dow, R., Barnsley, J., Tu, K., Jadad, A. R., & Lemieux-Charles, L. (2013). End-user support for primary care electronic medical records: a qualitative case study of users' needs, expectations, and realities. *Health Systems, 2*(3), 198–212. doi: 10.1057/hs.2013.6

Steis, M. R., Prabhu, V. V., Kang, Y., Bowles, K. H., Fick, D., & Evans, L. (2012). Detection of delirium in community-dwelling persons with dementia. *Online Journal of Nursing Informatics, 16*(1). Retrieved from http://ojni.org/issues/?=1274

van der Meijden, M. J., Tange, H. J., Troost, J., & Hasman, A. (2003). Determinants of success of clinical information systems: A literature review. *Journal of the American Medical Informatics Association, 10*(3), 235–243.

References Cited in Tables

Table 2.3

Ammenwerth, E., & de Keizer, N. (2004). An inventory of evaluation studies of information technology in health care. Trends in evaluation research, 1982-2002. *International Journal of Medical Informatics, 44*(1), 44–56.

Balas, A. E., Boren, S. A., & Griffing, G. (1998). Computerized management of diabetes: a synthesis of controlled trials. *Journal of the American Medical Informatics Association, 5*(S), 295–299.

Balas, A. E., Austin, S. M., Mitchell, J. A., Ewigman, B. G., Bopp, K. D., & Brown, G. D. (1996). The clinical value of computerized information services. *Archives of Family Medicine, 5*(5), 271–278.

Cramer, K., Hartling, L., Wiebe, N., Russell, K., Crumley, E., Pusic, M., & Klassen, T. P. (2003). *Computer-based delivery of health evidence: A systematic review of randomized controlled trials and systematic reviews of the effectiveness of the process of care and patient outcomes.* Final report, part I. Edmonton, AB: Heritage Foundation for Medical Research.

Delpierre, C., Cuzin, L., Fillaux, J., Alvarez, M., Massip, P., & Lang T. (2004). A systematic review of computer-based patient record systems and quality of care: More randomized clinical trials or a broader approach? *International Journal for Quality in Health Care, 15*(5), 407–416.

Garg, A. X., Adhikari, N. K. J., McDonald, H., Rosas-Arellano, M. P., Devereaux, P. J., Beyene, J., Sam, J., & Haynes, R. B. (2005). Effects of computer-based clinical decision support systems on practitioner performance and patient outcomes: A systematic review. *Journal of the American Medical Association, 293*(10), 1223–1238.

Kaushal, R., Shojania, K. G., & Bates, D. W. (2003). Effects of computerized physician order entry and clinical decision support systems on medication safety. *Archives of Internal Medicine, 163*(12), 1409–1416.

Kawamoto, K., Houlihan, C. A., Balas, E. A., & Lobach, D. F. (2005). Improving clinical practice using clinical decision support systems: A systematic review of trials to identify features critical to success. *British Medical Journal, 330*(7494), 765–768.

Mitchell, E., & Sullivan, F. A. (2001). A descriptive feast but an evaluative famine: systematic review of published articles on primary care computing during 1980-97. *British Medical Journal, 322*(7281), 279–282.

Montgomery, A. A., & Fahey, T. (1998). A systematic review of the use of computers in the management of hypertension. *Journal of Epidemiology and Community Health, 52*(8), 520–525.

Sullivan, F., & Mitchell, E. (1995). Has general practitioner computing made a difference to patient care? *British Medical Journal, 311*(7009), 848–852.

van der Loo, R. P., Gennip, E. M. S. J., Bakker, A. R., Hasman, A., & Rutten, F. F. H. (1995). Evaluation of automated information systems in healthcare: An approach to classify evaluative studies. *Computer Methods and Programs in Biomedicine, 48*(1/2), 45–52.

van der Meijden, M. J., Tange, H. J., Troost, J., & Hasman, A. (2003). Determinants of success of clinical information systems: A literature review. *Journal of the American Medical Informatics Association*, *10*(3), 235–243.

Walton, R., Dovey, S., Harvey, E., & Freemantle, N. (1999). Computer support for determining drug dose: systematic review and meta analysis. *British Medical Journal, 318*(7189), 984–990.

Table 2.5

British Columbia Ministry of Health [MOH]. (2011a). *Evaluating the benefits: Telehealth — TeleOncology*. Initial benefits evaluation assessment. Victoria, BC: Ministry of Health.

British Columbia Ministry of Health [MOH]. (2011b). *Evaluating the benefits: Telehealth — teleOphthalmology*. Inter Tribal Health Authority and the Ministry of Health Services. Victoria, BC: Ministry of Health.

British Columbia Ministry of Health [MOH]. (2013). *Panorama vaccine inventory module*. Benefits evaluation report. Victoria, BC: Ministry of Health.

Cousins, A., & Baldwin, A. (2014). *eChart ambulatory project: EMR benefits measurement in a tertiary care facility*. eHealth Conference, Vancouver, BC, June 3, 2014.

Deloitte. (2010). *National impacts of generation 2 drug information systems*. Technical Report, September 2010. Toronto: Canada Health Infoway. Retrieved from https://www.infoway-inforoute.ca/en/component/edocman/resources/reports/331-national-impact-of-generation-2-drug-information-systems-technical-report-full

Elliott, P., Martin, D., & Neville, D. (2014). Electronic clinical safety reporting system: A benefits evaluation. *Journal of Medical Internet Research*, *2*(1), e12.

Ernst & Young. (2014). *Connecting patients with providers: A pan-Canadian study on remote patient monitoring*. Toronto: Canada Health Infoway. Retrieved from https://www.infoway-inforoute.ca/en/component/edocman/resources/reports/benefits-evaluation/1890-connecting-patients-with-providers-a-pan-canadian-study-on-remote-patient-monitoring-executive-summary

Gartner, Inc. (2013). *British Columbia eHealth benefits estimates*. Stamford, CT: Author.

Hagens, S., Kwan, D., Savage, C., & Nenadovic, M. (2009a). The impact of diagnostic imaging investments on the Canadian healthcare system. *Electronic Healthcare, 7*(4), Online Exclusive.

Hagens, S., Kraetschmer, N., & Savage, C. (2009b). Findings from evaluations of the benefits of diagnostic imaging systems. In J. G. McDaniel (Ed.), *Advances in information technology and communication in health* (pp. 136–141). Amsterdam: IOS Press.

Lau, F., Lee, D., Quan, H., & Richards, C. (2010). An exploratory study to examine the use of SNOMED CT in a palliative care setting. *Electronic Healthcare, 9*(3), e12–e24.

Newfoundland and Labrador Centre for Health Information [NLCHI]. (2010). *Evaluating the benefits — Newfoundland and Labrador provincial telehealth program: Chronic disease management*. St. John's: Author.

Praxia Information Intelligence, & Gartner, Inc. (2010). *Telehealth benefits and adoption: Connecting people and providers across Canada*. Toronto and Stamford, CT: Authors.

PricewaterhouseCoopers [PWC]. (2013). *The emerging benefits of electronic medical record use in community-based care*. Toronto: Canada Health Infoway.

Table 2.6

Ahmadi, H., Rad, M. S., Nilashi, M., Ibrahim, O., & Almaee, A. (2013). Ranking the micro level critical factors of electronic medical records adoption using toopsis method. *Health Informatics — An International Journal (HIIJ), 2*(4), 19–32.

Cusack, C. M., Hook, J. M., McGowan, J., Poon, E., & Atif, Z. (2010). *Evaluation toolkit. Health information exchange projects, 2009 update* (AHRQ Publication No. 10-0056-EF). Rockville, MD: US Department of Health and Human Services.

Gu, Y., Warren, J., Day, K., Pollock, M., & White, S. (2012). Achieving acceptable structured eReferral forms. In K. Butler-Henderson & K. Gray (Eds.), *Proceedings of the 5th Australasian workshop on Health Informatics and Knowledge Management (HIKM, 2012), Melbourne* (pp. 31–40). Darlinghurst, Australia: Australian Computer Society.

Hotchkiss, D., Dianna, M., & Foreit, K. (2012). *How can routine health information systems improve health systems functioning in low-resource settings?* Chapel Hill, NC: MEASURE Evaluation.

Nguyen, L., & Bakewell, L. (2011). Impact of nursing information systems on residential care provision in an aged care provider. In *ACIS 2011: Proceedings of the 22nd Australasian Conference on Information Systems : Identifying the information systems discipline* (pp. 1–10). Sydney: AIS eLibrary.

Warren, J., Pollock, M., Day, K., Gu, Y., & White, S. (2011). *A framework for health IT evaluation*. Wellington, New Zealand: Ministry of Health.

Warren, J., Gu, Y., Day, K., & Pollock, M. (2012). Approach to health innovation projects: Learnings from eReferrals. *Health Care and Informatics Review Online, 16*(2), 17–23.

Table 2.7

British Columbia Ministry of Health [MOH]. (2011). *Evaluating the benefits: Telehealth — teleOphthalmology*. Inter Tribal Health Authority and the Ministry of Health Services. Victoria, BC: Ministry of Health.

British Columbia Ministry of Health [MOH]. (2013). *Panorama vaccine inventory module* (Benefits Evaluation Report). Victoria, BC: Ministry of Health.

Cousins, A., & Baldwin, A. (2014). *eChart ambulatory project: EMR benefits measurement in a tertiary care facility*. eHealth Conference, Vancouver, BC, June 3, 2014.

Eapen, B. R., & Chapman, B. (2015). Mobile access to ClinicalConnect: A user feedback survey on usability, productivity, and quality. *Journal of Medical Internet Research, 3*(2), e35. doi: 10.2196/mhealth.4011

Kuhn, K., & Lau, F. (2014). Evaluation of a shared electronic health record. *Healthcare Quarterly, 17*(1), 30–35.

Prince Edward Island [P.E.I.]. (2010). *Prince Edward Island drug information system — Evaluation report*. Charlottetown: Government of Prince Edward Island.

Chapter 3

Clinical Adoption Framework

Francis Lau, Morgan Price

3.1 Introduction

In 2006, Canada Health Infoway published the Benefits Evaluation (BE) Framework that was adapted from the Information System (IS) Success Model by DeLone and McLean (as cited in Lau, Hagens, & Muttitt, 2007). The BE Framework provides a conceptual model for understanding the quality, use and net benefits of eHealth adoption in healthcare organizations. The BE Framework has been well received by the healthcare community because it "made sense" as an organizing scheme when describing eHealth adoption and evaluation. However, the original IS Success Model was based on a stable business IS environment and did not take into account the organizational and social contexts. In 2009, we extended the BE Framework by incorporating a set of meso- and macro-level factors that could influence the success of eHealth systems (Lau, 2009). The extensions have led to the Clinical Adoption (CA) Framework described here.

This chapter describes the conceptual foundations of the CA Framework and the micro, meso and macro dimensions that made up this framework. We then describe the validation and use of this framework, and its implications on eHealth evaluation for healthcare organizations.

3.2 Conceptual Foundations

The CA Framework is built on theories and models from the disciplines of information systems, organization science, and health informatics. They include: the Information Technology Interaction Model by Silver, Markus, and Beath (1995); the Unified Theory of Acceptance and Use of Technology Model by Venkatesh, Morris, Davis, and Davis (2003); earlier work in implementation re-

search by Cooper and Zmud (1990); task-technology fit by Goodhue and Thompson (1995) and Ammenwerth, Iller, and Mahler (2006); managing change and risks by Kotter and Schlesinger (1979) and Paré, Sicotte, Jaana, and Girouard (2008); and the people and socio-organizational aspects of eHealth by Berg, Aarts, and van der Lei (2003), Kaplan, Brennan, Dowling, Friedman, and Peel (2001), Kaplan and Shaw (2004), and Stead and Lorenzi (1999). These published sources are described below.

3.2.1 Information Technology Interaction Model

The Information Technology Interaction Model, or ITIM, was introduced by Silver, Markus, and Beath in 1995 as a teaching model for Master of Business Administration (MBA) students. The model describes the effects of an information system interacting on an organization over time. There are four interrelated dimensions in ITIM: the information system, implementation process, organizational context, and the system's effects (Figure 3.1). Each of these dimensions is represented by a set of components and subcomponents, which are summarized below.

- Information system – functionality, interface, restrictiveness, guidance, and decision-making
- Implementation process – initiation, build/buy, introduction, and adaptation
- Organizational context – firm's structure, processes, strategies, culture, IT infrastructure, and external environment, more specifically:
  - Structure – de/centralization, functional/divisional/network, reporting relationships
  - Processes – order fulfillment, materials acquisition, product development
  - Strategies – differentiation, low-cost production, quality/service, right-sizing, just-in-time
  - Culture – artefacts, shared values, assumptions, individuality/teamwork, risk handling
  - IT infrastructure – hardware, software, databases, networks, training, personnel, skills
  - External environment – industry structure, competition, buyer/seller power, growth
- System's effects – use, consequences and adaptations, more specifically:
  - Use – whether the system is used or not, how it is used, by whom, and for what purpose

- Consequences – performance effects such as profits, effects on people such as power and role, and future flexibility for the organization
- Adaptations – feedback effects on the organization from performance, people, and flexibility

Since its publication in 1995, the ITIM has been applied and cited in many studies related to IS. One application is to use the ITIM's organization, implementation and effect dimensions as a conceptual scheme to critique, refine and develop additional Information Technology (IT) or IS theories and models. For instance, in his re-specification of DeLone and McLean's IS Success Model, Seddon (1997) argued the ITIM system's effects on use and consequences are similar to the DeLone and McLean model's net benefits, and that the greater IS use implied more consequences. Kohli and Limayen (2006) and Tams (2011) applied the ITIM as a foundational model to justify the legitimacy of IS as a reference discipline through its theoretical and methodological contributions in the areas of IS development, implementation, innovation, and business value. In healthcare, Ben-Zion, Pliskin, and Fink (2014) applied the ITIM dimensions in a literature review and prescriptive analysis to identify a set of critical success factors for the adoption of EHR systems.

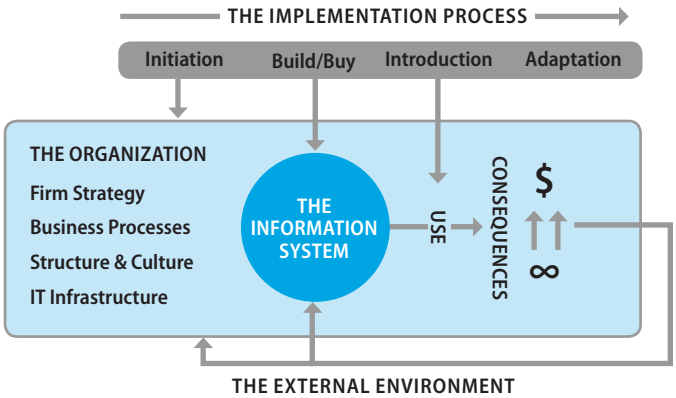


Figure 3.1. IT interaction model.

Note. From "The information technology interaction model: A foundation for the MBA core course," by M. S. Silver, M. L. Markus, and C. M. Beath, 1995, *Management Information Systems Quarterly, 19*(3), p. 366. Copyright 1995 by Regents of the University of Minnesota. Reprinted with permission.

3.2.2 Technology Acceptance Models

The original Technology Acceptance Model (TAM) by Davis (1989) and its variants (e.g., TAM2) published over the years are considered the most widely applied theory on an individual's acceptance of technology (Lee, Kozar, & Larsen, 2001; Yarbrough & Smith, 2007). In 2003, Venkatesh et al. published the Unified Theory

of Acceptance and Use of Technology (UTAUT) Model based on a synthesis of eight TAM-related models. The UTAUT combined the best features from these models and has emerged as one of the most widely cited models on technology acceptance. The UTAUT has four attributes that are considered the direct determinants of technology use intention and/or behaviour: performance expectancy, effort expectancy, social influence, and facilitating conditions (i.e., the perceived technical and organizational infrastructure in place to support IS use). There are also four other attributes that have a moderating effect on the direct determinants with respect to their influence on technology use intention and/or behaviour: gender, age, voluntariness, and experience. The UTAUT Model is shown in Figure 3.2.

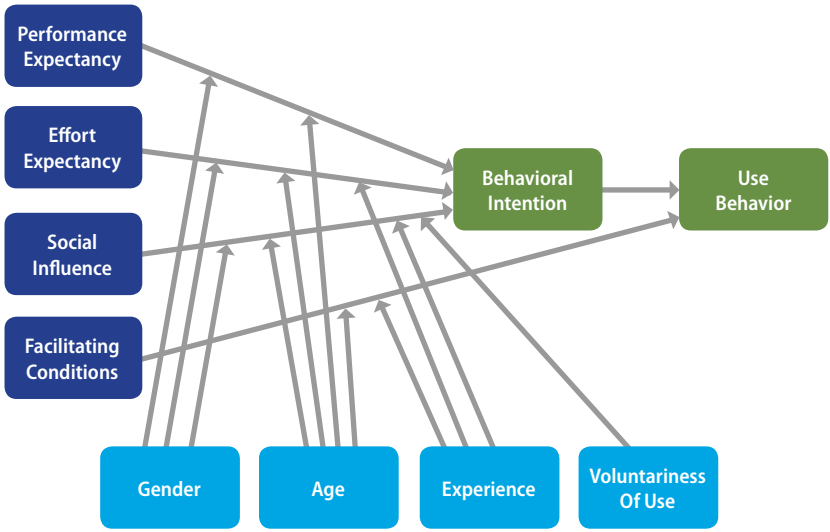


Figure 3.2. Unified theory of acceptance and use of technology.

Note. From "User acceptance of information technology: Toward a unified view," by V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, 2003, *Management Information Systems Quarterly, 27*(3), p. 447. Copyright 2003 by Regents of the University of Minnesota. Reprinted with permission.

Since its publication, the UTAUT Model has been applied in different healthcare settings to determine the acceptance of eHealth systems by care providers. For example, survey-based studies have examined the key organizational characteristics for successful telemedicine programs (Whitten, Holtz, & Nguyen, 2010), the factors that influence user acceptance of a hospital picture archiving and communication system (Duyck et al., 2008), acceptance of EMR systems by nurses, physician assistants, and nurse practitioners at the state level (Wills, El-Gayar, & Bennett, 2008), and perceptions of two outpatient electronic prescribing systems for primary care (Wang et al., 2009). Thus far, the UTAUT Model and its survey instrument have proved to be robust, valid and reliable when used in healthcare settings.

3.2.3 Implementation Research and Managing Change

There has been a significant amount of work done in IS implementation research regarding the theories, methods, processes and implications of IS implementation in organizations (e.g., Kukafka, Johnson, Linfante, & Allegrante, 2003). Examples are the technological diffusion approach by Cooper and Zmud (1990) and the improvisational model for change by Orlikowski and Hofman (1997). Of particular interest is the work on task-technology fit by Goodhue and Thompson (1995) and Ammenwerth et al. (2006) that focused on the relationships between an individual's performance and his or her technology-enabled work. The importance of managing organizational change and its effects on IS implementation has also been recognized (e.g., Lorenzi, 2000; Iles & Sutherland, 2001).

The organizational change model by Kotter (2007) and the project risk assessment framework by Paré et al. (2008) are examples of practice-based change management approaches applied to ensure successful IS implementation. To transform an organization, Kotter emphasized the need for a sense of urgency, a powerful guiding coalition, a communicated vision empowering those to act on the vision, focusing on short-term wins, consolidating improvement to produce more change, and institutionalizing the new approach. Similarly, Paré and colleagues offered a systematic approach to ensuring successful IS implementation by reducing risks along the technological, human, usability, managerial, strategic, and political dimensions.

3.2.4 People and Socio-organizational Aspects

In health informatics there has been a shift from a technical focus on the deployment of local eHealth systems to a broader focus of sociotechnical systems with the emphasis on people, organizational and social issues. In 1999, Stead and Lorenzi (1999) suggested the health informatics agenda should "acknowledge the foundation provided by the health system … the role of financial issues, system impediments, policy and knowledge in effecting change" (p. 341). Similarly, Kaplan and colleagues (2001) outlined an informatics research agenda that involved the use of different social inquiry methods depending on settings at the individual, institutional, trans-organizational and transnational levels. Kaplan and Shaw (2004) further outlined the directions for informatics evaluation to include the reshaping of institutional boundaries, changing work practices and standards, the politicization of healthcare, and changing roles for providers and consumers. The sociotechnical approaches advocated by Berg et al. (2003) also emphasized the social nature of healthcare work that can influence the success of eHealth systems, including meso- and macro-level processes such as the financial status of the organization, jurisdictional healthcare policy, and politics at both the institutional and national levels.

3.3 CA Framework Dimensions

The CA Framework has three conceptual views of eHealth adoption by clinicians in different settings (Lau, Price, & Keshavjee, 2011). These are the micro-, meso- and macro-level views of clinical adoption. They are described below.

- The *micro level* addresses the quality of the information, system and service associated with an eHealth system, its use and user satisfaction, and net benefits in terms of care quality, productivity and access. These are the same dimensions and categories that are defined in the BE Framework.

- The *meso level* addresses the people, organization and implementation dimensions that have a direct effect on the micro level eHealth adoption by clinicians. The people dimension is drawn from the constructs in the UTAUT, while the organization and implementation dimensions are from the ITIM, implementation research, and change management models described earlier.

- The *macro level* addresses healthcare governance, standards, funding, and societal trends as the environmental factors that have direct influence on the extent to which the meso level can affect clinical adoption at the micro level. These macro-level factors are based on the sociotechnical approaches that transcend organizations to include overall societal trends.

- At each level there is a *feedback loop* where the adoption efforts and results can reshape the higher levels. The CA Framework is shown in Figure 3.3 and the three views are elaborated next. The CA categories, subcategories and measures are summarized in the Appendix following the References section.

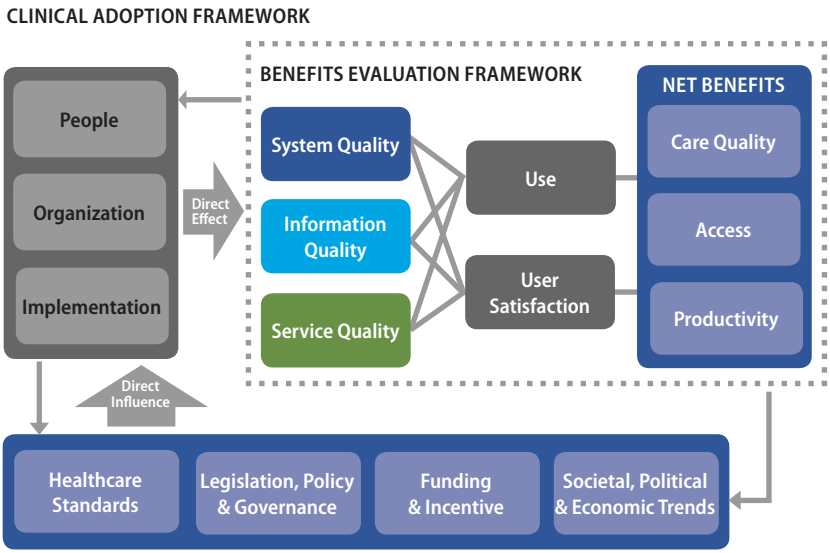


Figure 3.3. Clinical adoption framework with its micro, meso and macro dimensions.

Note. From "From benefits evaluation to clinical adoption: Making sense of health information system success in Canada," by F. Lau, M. Price, and K. Keshavjee, 2011, *Healthcare Quarterly, 14*(1), p. 41. Copyright 2011 by Longwoods™ Publishing Corp. Reprinted with permission.

3.3.1 Micro Level

At the micro level, our proposition is that successful clinical adoption of an eHealth system depends on its HIT quality, usage quality and net benefits. These are elaborated below.

- *HIT Quality* refers to the accuracy, completeness and availability of the clinical information content of an eHealth system; the features, performance and security of the system; and responsiveness of the system's support services.

- *Usage Quality* refers to eHealth system usage intention/pattern; and user satisfaction in terms of usefulness, ease-of-use and competency.

- *Net Benefits* refer to changes in care quality, access and productivity as a result of eHealth adoption by clinicians. Care quality covers patient safety, appropriateness/effectiveness and health outcomes. Access covers provider/patient participation and availability/access to services. Productivity covers care coordination, efficiency and net cost.

Our rationale is that the better the quality of the eHealth system adopted, the more it will be embraced by satisfied clinicians, leading to greater tangible net benefits over time.

3.3.2 Meso Level

At the meso level, our proposition is that successful clinical adoption depends on the people, organization and implementation process. These are elaborated below.

- *People* refers to all types of individuals or groups in the healthcare system having to do with eHealth in some way, their personal characteristics and expectations, as well as their roles and responsibilities within the eHealth system.

- *Organization* refers to how the system fits with the organization's strategy, culture, structure/processes, information infrastructure and return on value.

- *Implementation* refers to the eHealth adoption stages, project management approaches, and the extent of eHealth-practice fit planned in the future and operating at present.

Our rationale is that higher eHealth adoption can occur in the organization if clinicians have experience and clear expectations in using the system. Moreover, the system will be seen as adding value if it is designed to support organizational performance goals. To do so, the implementation process must be carefully planned, executed and managed throughout its life cycle. This ensures the eHealth system fits into the day-to-day work practices of clinicians. When these meso-level factors are aligned with those at the micro level, we can expect further magnified improvements in eHealth system quality, usage and net benefits.

3.3.3 Macro Level

At the macro perspective, our proposition is that successful clinical adoption depends on the environmental contexts with respect to governance, standards, funding and trends. These are elaborated below.

- *Governance* refers to the influence of governing bodies, legislative acts, and the regulations or policies covering such bodies as professional associations/colleges, advocacy groups and their attitudes toward eHealth.

- *Standards* refer to the types of eHealth, organizational performance and professional practice standards in place.

- *Funding* refers to the payment, remuneration, and incentive programs in place.

- *Trends* refer to public expectations, and the overall socio-political and economic climates toward technologies, eHealth and health care as a whole.

Our rationale is that higher eHealth adoption by clinicians can be achieved if the organization aligns its effort with the macro environmental factors that influence clinical adoption. For instance, organizations should embrace eHealth systems that conform to industry-wide interoperable standards, help achieve external performance targets, and adapt to the changing scope of professional practice in care delivery. Where feasible, organizations should take advantage of incentives that encourage clinical adoption such as subsidized eHealth system deployment and automated patient safety surveillance. Adhering to established health information protection legislations, policies and practices with strong governance involving multiple stakeholders can further enhance clinical adoption through trust and relationship building. Lastly, staying abreast of the socio-political and economic trends — such as encouraging citizens to better manage their own health through the use of personal health records — allows the organization to be proactive in its eHealth planning and deployment efforts.

3.4 CA Framework Usage

3.4.1 Validation of the CA Framework

The CA Framework underwent three validation steps when it was introduced. First was a comparison of the framework elements (i.e., dimensions, categories and measures) against those identified in a meta-review of eHealth evaluation systematic reviews (Lau, Price, Kuziemsky, & Gardner, 2010). Second was a consultation session with Canadian eHealth practitioners to determine if they agreed with the framework elements (Lau & Charlebois, 2009). Third was a comparison against the questions/measures used in survey instruments of published eHealth adoption and evaluation studies (Oh, 2009). The three steps are summarized below.

- In a *meta-review* of 50 eHealth evaluation systematic reviews published between 1995 and 2008, Lau et al. (2010) were able to map most of the evaluation measures from the reviews to the micro-level dimensions of the CA Framework. They also identified measures that did not fit the micro level and created new categories for them which were patient/provider, implementation, incentive, policy/legislation, change improvement, and interoperability. These factors mapped nicely under the meso- and macro-level dimensions of the CA Framework.

- In 2009 Infoway held a *consultation session* with 23 eHealth practitioners from across Canada that provided their anonymized written feedback on the CA Framework. The practitioners responded to questions on whether the framework made sense, whether concepts were missing or required revisions, as well as their interest in, and the effort needed to apply the framework in their organizations. Based on their feedback, revisions were made to streamline the framework, for example by dropping the network dimension and making the people dimension more prominent (Lau & Charlebois, 2009).

- Oh (2009) compared the CA Framework elements against 16 *published survey instruments*. They included 13 instruments from the Health IT Survey Compendium section of the Agency for Healthcare Research & Quality (AHRQ) Health IT website (AHRQ, 2010) and three from Canada Health Infoway. Of the 16 instruments examined, only the Infoway System and Use Assessment Survey items mapped to all 20 micro-level elements. At the meso level the 16 instruments mapped between 0 and 11/12 of the elements. At the macro level they mapped poorly from 0 to 5/12 elements. No question items were found missing from the framework, which suggested it was sufficiently comprehensive for all areas of eHealth.

3.4.2 Use of the CA Framework

The CA Framework provides an overarching conceptual model that makes sense of eHealth adoption by clinicians. Healthcare organizations involved with eHealth adoption should address as needed the micro-, meso- and macro-level factors described in this framework to achieve eHealth success. Given the large number of factors that affect clinical adoption, an organization should focus on a subset of these factors when evaluating its eHealth adoption effort and impacts. To apply the CA Framework, one needs different methods and tools to evaluate whether the factors are associated with the extent of adoption and impacts desired and/or achieved. Examples of evaluation methods that can be applied before, during and after adoption of an eHealth system are the Infoway System and Use Assessment (SUA) survey and the Rapid Response Evaluation Methods (RREM) from the eHealth Observatory (Lau, 2010). The RREM is made up of a suite of evaluation tools for conducting usability, workflow, system/data quality and impact studies, and practice reflections for different implementation stages. Depending on need, other evaluation methods can be applied to examine particular aspects of clinical adoption in specific settings.

To illustrate, an organization in the process of implementing a picture archiving and communication system (PACS) may wish to focus on specific micro-level factors in the CA Framework by examining the extent to which the quality of the

PACS, its perceived usefulness, and actual system usage can affect the productivity of the clinicians and their workflow coordination. By conducting the SUA survey and RREM workflow analysis before and after PACS deployment, one can compare the extent of work practice change brought on by the system. On the other hand, an organization with a suite of existing eHealth systems such as order entry or lab and pharmacy systems may focus on specific meso-level people and organization factors to improve their clinical adoption. By conducting the RREM impact assessment surveys, one can identify areas that require attention such as the extent of eHealth alignment with the organization's strategy, technical infrastructures and clinician expectations. Lastly, a jurisdiction wishing to evaluate the success of its primary healthcare EMR strategy may apply the RREM reactive analysis to see if the macro-level factors are adequately addressed. These may include EMR alignment with industry-wide eHealth standards, professional practice scope, medical service fee schedule, privacy legislations for patient record exchange, and societal expectations of value for money in EMR investments.

Since its debut in 2011, the CA Framework has been applied, adapted and cited in over 30 studies and publications. Examples where the CA Framework was applied are the ambulatory care clinic EMR evaluation study in a British Columbia health region by Lau, Partridge, Randhawa, and Bowen (2013) and a fuzzy modelling study to identify key meso-level factors for successful EMR adoption in eight Malaysian primary care clinics (Ahmadi et al., 2013). There are also two literature reviews where the CA Framework was applied as a conceptual scheme to organize the review findings (Lau, Price, Boyd, Partridge, Bell, & Raworth, 2012; Bassi, Lau, & Lesperance, 2012). In a coordinated Canadian EHR strategy white paper, Lau, Price, and Bassi (2014) adapted the CA Framework as a new eHealth Value Framework by expanding the investment, value and lag time aspects of eHealth adoption. In Finland, the National Institute for Health and Welfare incorporated the meso- and macro-level dimensions of the CA Framework into its eHealth Evaluation Framework to assess health information system implementation at the national level (Hypponen et al., 2011). See Table 3.1 for examples of studies where the CA Framework has been applied.

The CA Framework has been cited in different publications related to eHealth strategy, adoption and evaluation by health informaticians in several countries. For example, Axelsson and Melin (2014) acknowledged the importance of context when identifying critical success factors in Swedish eHealth systems. Yusof, Khodambashi, and Mokhtar (2012) cited the need to consider HIT-practice fit (part of the meso dimension in the CA Framework) as part of their lean method to study the implementation of a critical care information system in Malaysia. Similarly, Viitanen and colleagues (2011) emphasized the need to examine the contextual aspect of usability (i.e., eHealth-practice fit) when evaluating Finnish clinical IT systems. In their study of clinical governance and EMR adoption in the Australian primary care setting, Pearce, de Lusignan, Phillips, Hall, and

Table 3.1
Canadian Evaluation Studies where the CA Framework was Applied

Authors	Setting	eHealth system	Evaluation Focus	Design/Methods	Indicators/Measures	Results
Ahmadi et al. (2013) Malaysia	Eight primary care clinics	EMR systems	Identification of most influential meso-level factors – people, organization, implementation	Survey, modelling with fuzzy technique for order performance by similarity to ideal solution (TOPSIS), analytical hierarchy process (AHP)	Likert-scale surveys with 16 parameters under meso level – people, organization and implementation	Influential factors found were time investment, screen/room, hybrid system, planning, resource training, workflow and value
Bassi et al. (2012)	Physician offices	EMR systems	Perceived impact from surveys	Systematic review of published surveys, impact factors mapped to CA Framework, meta-analysis of selected impact areas	Seven impact areas with standardized positive-negative-mixed views by user/non-user	Mostly positive views regardless of user status, area with mostly mixed views is security and privacy
Hypponen et al. (2011)	All settings	Health information systems	Large-scale lessons of eHealth system implementation	Literature review, framework design and physician surveys	Dimensions, categories, measures of eHealth success	Evidence categories for eHealth success with baseline results
Lau et al. (2012)	Physician offices	EMR systems	Impacts, success factors and lessons	Systematic review of primary studies on EMR impact, organized by CA Framework	Six impact areas with proportions of positive-negative-neutral studies, factors that influence success, and common lessons	51% studies positive, 19% negative and 30% neutral; 48 factors influenced success. Five repeated lessons

Table 3.1

Canadian Evaluation Studies where the CA Framework was Applied

Authors	Setting	eHealth system	Evaluation Focus	Design/Methods	Indicators/Measures	Results
Lau et al. (2013)	Ambulatory care clinic in a health region	Ambulatory EMR system	Post-implementation formative evaluation of EMR impact based on CA Framework	Rapid evaluation methods with surveys, interviews, usability/workflow analysis, project risk assessment, data quality and document review, group reflections	EMR quality, use and satisfaction; care coordination and efficiency; people roles, expectations and experiences; organization process strategy and infrastructure; implementation process and EMR-practice fit	Micro- and meso-level issues affected EMR adoption, some perceived benefits reported in care coordination and efficiency, challenges and lessons identified
Lau et al. (2014)	Canada-wide	Any eHealth system	A coordinated EHR strategy based on the CA Framework	Literature reviews on Canadian and international evaluation studies	Investment, adoption, lag time and value dimensions with suggested measures	A coordinated EHR strategy with 10 implementation steps

Traveglia (2013) identified similar meso- and macro-level factors from the CA Framework that influenced EMR acceptance.

The CA Framework has also been cited in a number of graduate student theses related to eHealth. Examples include the study of EMR data quality and payment incentives in primary care (Bowen, 2013), the meaningful use in primary care EMRs (Watt, 2014), a review of health information exchanges' success factors (Ng, 2012), an evaluation of multidisciplinary cancer care conference platforms (Ghaznavi, 2012), end user support for primary care EMRs (Dow, 2012), and critical success factors for Malaysian public hospital information systems (Abdullah, 2013).

3.5 Implications

The current CA Framework requires further work to improve its validity, relevance and utility. Some of the meso- and macro-level factors in the framework need to be refined as specific measures that can be applied and quantified in field settings. In particular, evaluation methods that measure specific factors in the CA Framework are needed in order for it to be applied more widely across different types of eHealth systems and organizational settings. Additional methods and tools are also required to evaluate factors that are not currently addressed, especially in the areas of health outcomes at the micro level, return on value at the meso level, and governance, funding and standards at the macro level.

Despite the limitations, it is important to keep in mind that to make major strides forward with clinical adoption of eHealth systems, healthcare organizations need to share a common vision of what constitutes eHealth success. The CA Framework provides a common ground by which eHealth adoption by clinicians can be described, measured, compared and aggregated as empirical evidence over time.

3.6 Summary

This chapter described the CA Framework for determining eHealth success. It is an extension of the BE Framework that takes into account the contextual factors involved. The CA Framework has three conceptual dimensions at the micro, meso and macro levels. Each dimension has its own set of factors that define eHealth success. The CA Framework has undergone an initial validation, and has been proposed as an overarching framework to plan, conduct and report eHealth evaluation studies. The advantage of having a common evaluation framework is the ability to measure, compare and aggregate eHealth evidence in a consistent manner across different eHealth systems and healthcare settings.

References

Abdullah, Z. S. (2013). *Hospital information systems implementation framework: Critical success factors for Malaysian public hospitals* (Doctoral thesis, Curtin University, Perth Australia). Retrieved from http://espace.library.curtin.edu.au/R?func=dbin-jump-full&local_base=gen01-era02&object_id=192723.

Agency for Healthcare Research and Quality (AHRQ). (2010). *Health IT survey compendium.* Rockville, MD: Author. Retrieved from https://healthit.ahrq.gov/health-it-tools-and-resources/health-it-survey-compendium

Ahmadi, H., Nilashi, M., Darvishi, M., Ibrahim, O., Zakaria, R., Zolghadri, H., & Alizadeh, M. (2013). Fuzzy multi-criteria approaches for evaluating the critical factors of electronic medical record adoption. *Review of Contemporary Business Research, 3*(2), 1–24.

Ammenwerth, E., Iller, C., & Mahler, C. (2006). IT-adoption and the interaction of task, technology and individuals: A fit framework and a case study. *BMC Medical Informatics and Decision Making, 6*(3), 1–13.

Axelsson, K., & Melin, U. (2014). Contextual factors influencing health information systems implementation in public sector: Investigating the explanatory power of critical success factors. In M. Janssen, H. J. Scholl, M. A. Wimmer, & F. Bannister (Eds.), *Proceedings of the 13th Conference on Electronic Governance, Dublin, Ireland, September 1 to 3, lecture notes in computer science* (pp. 59–71). Laxenburg, Austria: International Federation for Information Processing.

Bassi, J., Lau, F., & Lesperance, M. (2012). Perceived impact of electronic medical records in physician office practices: A review of survey-based research. *Interactive Journal of Medical Research, 1*(2), e3.1–e3.23.

Ben-Zion, R., Pliskin, N., & Fink, L. (2014). Critical success factors for adoption of electronic health record systems: Literature review and prescriptive analysis. *Information Systems Management, 31*(4), 296–312.

Berg, M., Aarts, J., & van der Lei, J. (2003). ICT in health care: sociotechnical approaches. *Methods of Information in Medicine, 42*(4), 287–301

Bowen, M. (2013). *Understanding, evaluating and enhancing EMR adoption in a primary care setting* (Unpublished master's thesis). University of Victoria, Victoria, BC, Canada.

Cooper, R. B., & Zmud, R. W. (1990). Information technology implementation research: A technological diffusion approach. *Management Science, 36*(2), 123–139.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *Management Information Systems Quarterly*, *13*(3), 319–339.

Dow, R. (2012). *The impact of end-user support on EMR success in Ontario primary care: A critical case study* (Unpublished master's thesis). University of Toronto, Toronto, ON, Canada.

Duyck, P., Pynoo, B., Devolder, P., Voet, T., Adang, L., & Vercruysse, J. (2008). User acceptance of a picture archiving and communication system. *Methods of Information in Medicine, 47*(2), 149–156.

Ghaznavi, F. (2012). *Design and evaluation of a multidisciplinary cancer conferencing platform* (Unpublished master's thesis). Chalmers University of Technology, Göteborg, Sweden.

Goodhue, D. L., & Thompson, R. L. (1995). Task-technology fit and individual performance. *Management Information Systems Quarterly, 19*(2), 213–236.

Hypponen, H., Viitanen, J., Reponen, J., Doupi, P., Jormanainen, V., Lääveri, T., … Hämäläinen, P. (2011). Large-scale eHealth systems: Providing information to support evidence-based management. In L. Van Gemert-Pijnen, H. C. Ossebaard, & P. Hämäläinen (Eds.), *ETELEMED: 3rd International Conference on eHealth, Telemedicine, and Social Medicine, February 23 to 28, Gosier, Guadeloupe, France* (pp. 89–95).

Iles, V., & Sutherland, K. (2001). *Organizational change: A review for healthcare managers, professionals and researchers.* London: National Coordinating Centre for National Health Service Delivery and Organization R&D.

Kaplan, B., Brennan, P. F., Dowling, A. F., Friedman, C. P., & Peel, V. (2001). Toward an informatics research agenda: Key people and organizational issues. *Journal of American Medical Informatics Association, 8*(3), 235–241.

Kaplan, B., & Shaw, N. T. (2004). Future directions in evaluation research: People, organizational and social issues. *Methods of Information in Medicine, 43*(3), 215–231.

Kohli, R., & Limayen, M. (2006). *Information systems research: Reference disciplines and theoretical contributions*. Proceedings of the Americas Conference on Information Systems (AMCIS), Acapulco, Mexico (Paper # 3).

Kotter, J. P., & Schlesinger, L. A. (1979). Choosing strategies for change. *Harvard Business Review, 57* (March-April), 106–114.

Kotter, J. P. (2007). Leading change: Why transformation efforts fail. *Harvard Business Review*, *85*(1), 96–103.

Kukafka, R., Johnson, S. B., Linfante, A., & Allegrante, J. P. (2003). Grounding a new information technology implementation framework in behavioural science: A systematic analysis of the literature on IT use. *Journal of Biomedical Informatics, 36*(3), 218–227.

Lau, F. (2009). Extending the Infoway benefits evaluation framework for health information systems. In J. G. McDaniel (Ed.), *Advances in information technology and communication in health* (pp. 406–413). Amsterdam: IOS Press.

Lau, F. (2010). EMR Evaluation Toolkit v3.0. Retrieved from http://ehealth.uvic.ca/resources/tools/tools.php

Lau, F., & Charlebois, M. M. (2009). *Validating the clinical adoption framework.* Clinical Adoption Forum, Nov 3, 2009 (unpublished).

Lau, F., Hagens, S., & Muttitt, S. (2007). A proposed benefits evaluation framework for health information systems in Canada. *Healthcare Quarterly, 10*(1), 112–118.

Lau, F., Partridge, C., Randhawa, G., & Bowen, M. (2013). Applying the clinical adoption framework to evaluate the impact of an ambulatory electronic medical record. *Studies in Health Technology and Informatics*, *83*, 15–20.

Lau, F., Price, M., & Keshavjee, K. (2011). From benefits evaluation to clinical adoption: Making sense of health information system success in Canada. *Healthcare Quarterly, 14*(1), 39–45.

Lau, F., Price, M., Kuziemsky, C., & Gardner, J. (2010). A review on reviews of health information system studies. *Journal of American Medical Informatics Association, 17*(6), 637–645.

Lau, F., Price, M., Boyd, J., Partridge, C., Bell, H., & Raworth, R. (2012). Impact of electronic medical record on physician practice in office settings: A systematic review. *BMC Medical Informatics and Decision Making, 12*(10), 1–10.

Lau, F., Price, M., & Bassi, J. (2014). *Toward a coordinated electronic health record strategy for Canada. Creating strategic change in Canadian healthcare.* Conference Series 2, Queen's University. Toronto, May 14-15.

Lee, Y., Kozar, K. A., & Larsen, K. R. T. (2001). The technology acceptance model: Past, present and future. *Communications of the Association for Information Systems, 12*, 752–780.

Lorenzi, N. (2000). Managing change: An overview. *Journal of American Medical Informatics Association 7*(2), 116–124.

Ng, Y. (2012). *Key characteristics of health information exchange: A scoping review* (Unpublished master's thesis). University of Victoria, Victoria, BC, Canada.

Oh, J. (2009). Clinical adoption framework — Survey mapping, December 4, 2009 (Unpublished).

Orlikowski, W. J., & Hofman, J. D. (1997). An improvisational model for change management: The case of groupware technologies. *Sloan Management Review, 38*(2), 11–21.

Paré, G., Sicotte, C., Jaana, M., & Girouard, D. (2008). *Prioritizing clinical information system project risk factors: a Delphi study.* Paper presented at the Proceedings of the 21st Hawaii International Conference on System Sciences, Hawaii, January 5 to 7.

Pearce, C. M., de Lusignan, S., Phillips, C., Hall, S., & Travaglia, J. (2013). The computerized medical record as a tool for clinical governance in Australian primary care. *Interactive Journal of Medical Research, 2*(2), e26.

Seddon, P. B. (1997). A respecification and extension of the DeLone and McLean model of IS success. *Information Systems Research, 8*(3), 240–253.

Silver, M. S., Markus, M. L., & Beath, C. M. (1995). The information technology interaction model: A foundation for the MBA core course. *Management Information Systems Quarterly, 19*(3), 361–390.

Stead, W., & Lorenzi, N. M. (1999). Health informatics: Linking investments to value. *Journal of American Medical Informatics Association, 6*(5), 341–348.

Tams, S. (2011). On the legitimacy of IS as an independent discipline: Research that sets the field apart. *Proceedings of the 14th Southern Association for Information Systems Conference* (pp. 170–174). Atlanta, GA, March 25-26.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *Management Information Systems Quarterly, 27*(3), 425–478.

Viitanen, J., Hypponen, H., Laaveri, T., Vanska, J., Reponen, J., & Winblad, I. (2011). National questionnaire study on clinical ICT systems proofs: Physicians suffer from poor usability. *International Journal of Medical Informatics, 80*(10), 708–725.

Watt, R. (2014). *Does customized in-practice support improve EMR meaningful use in primary care? Evidence from a retrospective mixed methods evaluation. MSc Project in Health Informatics.* University of Victoria, Victoria, BC.

Wang, C. J., Patel, M. H., Schueth, A. J., Bradley, M., Wu, S., Crosson, J. C., Glassman, P. A., & Bell, D. S. (2009). Perceptions of standards-based electronic prescribing systems as implemented in outpatient primary care: A physician survey. *Journal of American Medical Informatics Association, 16*(4), 493–502.

Whitten, P., Holtz, B., & Nguyen, L. (2010). Keys to a successful and sustainable telemedicine program. *International Journal of Technology Assessment in Health Care, 26*(2), 211–216.

Wills, M. J., El-Gayar, O. F., & Bennett, D. (2008). Examining healthcare professionals' acceptance of electronic medical records using UTAUT. *Issues in Information Systems, 9*(2), 396–401.

Yarbrough, A. K., & Smith, T. B. (2007). Technology acceptance among physicians: A new take on TAM. *Medical Care Research Review, 64*(6), 650–672.

Yusof, M. M., Khodambashi, S., & Mokhtar, A. M. (2012). Evaluation of the clinical process in a critical care information system using the Lean method: a case study. *BMC Medical Informatics and Decision Making, 12*(150), 1–14.

Appendix

CA Framework Dimensions, Categories and Definitions

Dimension	Category	Definitions of Suggested Measures
Micro Level		
HIS Quality	Information	Content – completeness, accuracy, relevance and comprehension
	System	Functionality – type and level of features available Performance – Accessibility, reliability and system response time Security – type and level of features available
	Service	The degree to which an individual believes HIS is important, can improve job performance and infrastructures exist to support its adoption
	Roles and Responsibilities	The position, function and obligation of an individual/group in relation to HIS adoption, for example, being a stakeholder, leader, champion and project sponsor
Use and User Satisfaction	Use	User behaviour and pattern – type, frequency, duration, location and flexibility of actual usage Self-reported use – type, frequency, duration, location and flexibility of perceived usage Intention to use – proportion of and reasons for current non-users to become users
	Satisfaction	The degree to which an individual's age, gender, education, experience and expertise can affect the adoption of HIS
Net Benefits	Care Quality	Patient safety – preventable errors, surveillance/monitoring, and risk/error reduction Appropriateness and effectiveness – adherence, compliance, practices, continuity of care Health outcomes – clinical outcomes and changes in health status from eHealth interventions
	Productivity	Efficiency – resource use, improvement in output, management, efficiency and capability Care coordination – care provision by team and continuity of care across continuum Net cost – monetary avoidance, reductions, actual/projected savings
	Access	Ability to access service – availability, diversity, timeliness and consolidation of services Patient/caregiver participation – self-management and access to own information

CA Framework Dimensions, Categories and Definitions

Dimension	Category	Definitions of Suggested Measures
Meso Level		
People	Individuals and Groups	Types of individuals/groups who can affect the adoption of HIS, including patients/clients and families, healthcare providers and managers, policy planners, and stakeholder groups
	Personal Characteristics	The degree to which an individual's age, gender, education, experience and expertise can affect the adoption of HIS
	Personal Expectations	The degree to which an individual believes HIS is important, can improve job performance and infrastructures exist to support its adoption
	Roles and Responsibilities	The position, function and obligation of an individual/group in relation to HIS adoption, for example, being a stakeholder, leader, champion and project sponsor
Organization	Strategy	A set of coordinated activities designed to achieve the overall mandate and objectives of the organization, including HIS adoption
	Culture	The ingrained set of shared values, beliefs and assumptions acquired by members of an organization over time, including their views toward HIS
	Structure and Processes	Organizational functioning, including governance, configuration, reporting relationships, communication, as well as business and patient care processes such as continuity of care
	Info and Infrastructure	HIS governance/management, technical architectures, information assets, level of integration and privacy/security in place or planned
	Return on Value	Economic return on HIS investment in terms of cost benefit, effectiveness, utility and avoidance; business case, return on investment, value propositions, benefits realization
Implementation	Stage	HIS adoption stages from initiation, build/buy, introduction to adaptation
	Project	The planning, activities and resources for HIS adoption, including scope, objectives, constraints, targets, governance, methodology, commitment, communication, training, risks, monitoring, reporting and expectations
	HIS-Practice Fit	The degree of fit between the HIS and organizational work practices, and the extent of change from HIS adoption

CA Framework Dimensions, Categories and Definitions

Dimension	Category	Definitions of Suggested Measures
Macro Level		
Governance	Legislative Acts	The types of HIS related legislative acts, such as health information and privacy laws that govern the adoption of HIS
	Regulations and Policies	The types of HIS related regulations/policies, such as data access and security/privacy guidelines
	Governance Bodies	The types of accountability and decision making structures in place regarding the adoption of HIS
Standards	HIS Standards	The types of data, messaging, terminology and technology standards that influence the healthcare industry as a whole with respect to HIS adoption
	Performance Standards	The types of organizational performance standards in place such as those for accreditation of healthcare facilities and performance targets
	Practice Standards	The desired level of professional competency, knowledge, skills and performance in the workplace, including HIS adoption
Funding	Remunerations	The types of compensation available, such as alternative payment schemes to entice change at the individual, practice and organizational levels
	Added Values	General expectations on the return-on-value from the adoption of HIS such as improved patient safety and access to care
	Incentive Programs	The types of reward programs available that entice change at the individual, practice and organizational levels
Trends	Societal Trends	The general expectations of the public toward healthcare and HIS
	Political Trends	The general political climates toward healthcare and HIS
	Economic Trends	The general economic investment climates toward healthcare and HIS

Chapter 4

Clinical Adoption Meta-Model

Morgan Price

4.1 Introduction

The Clinical Adoption Meta-Model (CAMM) was developed to support those implementing, studying and evaluating health information systems (HIS) when they are planning evaluations of HIS deployments, and how they are used and incorporated into practice over time (Price & Lau, 2014). This model can inform expectations of stakeholders and evaluation plans so that the correct types of evaluation metrics are considered at appropriate times after an HIS implementation. The CAMM was designed to be accessible to evaluators and stakeholders.

This chapter will begin with conceptual foundations; it briefly describes several common adoption models (some of which are found elsewhere in this handbook). It will outline the four *dimensions* of the CAMM and then illustrate several *CAMM archetypes* or representative adoption trajectories. The archetypes are followed by a real-world illustration of how the CAMM can guide a benefits evaluation plan.

4.2 Conceptual Foundations

There are several general adoption models that have been developed to inform adoption such as the Technology Acceptance Model or TAM (Lee, Kozar, & Larsen, 2003) and TAM 2 (Holden & Karsh, 2010), the Unified Theory of Acceptance and Use of Technology or UTAUT (Venkatesh, Morris, Davis, & Davis, 2003), the IS success model (Delone & McLean, 2003), and the diffusion of innovation (Rogers & Shoemaker, 1971) to name a few. Many of these have been applied to describe or explain adoption of HIS and other health technologies, such as the TAM (Holden & Karsh, 2010) and diffusion of innovation (Greenhalgh, Robert, Macfarlane, Bate, & Kyriakidou, 2004).

Several adoption models have been developed for specific types of HIS. Healthcare Information and Management Systems Society (HIMSS) Analytics has three Electronic Medical Record (EMR) adoption models for U.S. hospitals, Canadian hospitals, and U.S. ambulatory EMRs (Palacio, Harrison, & Garets, 2010; Pettit, 2013). The picture archiving and communication system (PACS) maturity model (van de Wetering & Batenburg, 2009; van de Wetering, Batenburg, & Lederman, 2010) describes functionality and integration of PACS systems into hospital workflows. The EMR adoption model (Price, Lau, & Lai, 2011) assesses the use of office-based EMRs over ten functional categories to describe current adoption of the EMR in practice, similar to HIMSS.

In HIS adoption evaluation, we are interested in understanding how health information systems are adopted into healthcare in meaningful ways that improve patient outcomes, quality and sustainability of the healthcare system (Wu, Chaudhry, Wang, & Maglione, 2006). Without understanding the adoption process, we may make inaccurate assumptions about the HIS and attribute the HIS to benefits or lack of benefits seen in evaluation.

4.3 The Four Dimensions of the CAMM

The CAMM was developed to help consider and describe adoption post-deployment of an HIS across four dimensions over time. Figure 4.1 illustrates the CAMM with its four dimensions: *availability*, *use*, *behaviour*, and *outcomes*. The four dimensions are dependent on each other (e.g., use requires availability) and should be considered collectively when planning an evaluation. The CAMM intentionally focuses on the four dimensions to help shape a focused understanding of adoption over time.

The CAMM was designed to apply to a range of health information systems. Thus, the specific measures and metrics in each of the dimensions would depend on the specific HIS or HIS component being deployed and evaluated. Also, the timelines will vary with the specific HIS being evaluated, how it is being deployed, and the context into which it is being deployed. Smaller components and apps may be quickly adopted and show early outcome changes in shorter periods of time than larger, more comprehensive systems or wider deployments that may take years to adopt.

Changes in the four dimensions are dependent on many factors beyond just time, such as: the HIS itself, its deployment plan, training of users, user expectations, IT support, related information systems, culture, funding, and organizational and jurisdictional regulations. When considering metrics and seeking to understand successes or failures within and across dimensions, it is important to consider, broadly, the factors that can influence adoption. The same HIS may (and likely will) have markedly different adoption trajectories depending on where and how it is deployed.

4.3.1 CAMM Dimension: Availability

Availability is the first dimension. Availability is defined as the end user's ability to interact with the HIS and its content, when and where needed. Availability can have multiple aspects and here we consider three: user access, system availability, and content availability. *User access* is the ability for end users to access the system. This can be measured with, for example, the number of user accounts, the numbers of users trained, or the number of accounts with remote access. *System availability* describes how available the HIS is to its intended end users. This can be measured with, for example, metrics of HIS uptime, the number of terminals deployed, or platforms supported. *Content availability* considers the information that is accessible in or through the HIS. Content could include patient health data (e.g., lab results) or knowledge base information (e.g., drug monograms, rules for decision support). Content availability can be considered in terms of breadth (types of content), depth (amount of each type), and currency (how quickly the content is updated and available). As availability increases, one would expect the potential to use would increase. An HIS that has only a few trained users or that is only turned on for a few hours a day or that lacks content may not be used extensively.

4.3.2 CAMM Dimension: Use

Use is the second dimension of the CAMM and describes the actual interactions of the intended end users with the HIS. Use is dependent on availability and has two aspects: *use* of the system and *user experience*. *Use* can be measured through a number of metrics, such as: number of log-ins, duration of time the system is used, locations from which the HIS is used, areas of the HIS that are used. *User experience* describes the subjective experience of end users when using the system. User experience should consider the user's internal state and the context of the interaction (Hassenzahl & Tractinsky, 2006). Intention to use is excluded from the CAMM as this model specifically describes actual adoption and, thus, use and the user experience of that use are considered, not intention to use; intention to use is included in some other models (see chapter 2, for example). Intention could be considered in pre-deployment evaluations or could be considered when understanding why a system was not used.

4.3.3 CAMM Dimension: Clinical (Health) Behaviour

Behaviour is the third dimension of the CAMM. It describes meaningful adaptation of clinical or health workflows to leverage the HIS features. Behaviour can be considered in terms of two aspects: *general capacity* and *specific behaviours*. General capacity is a global change in the healthcare organization. General capacity measures could include the increase or decrease in the number of patients seen per day, or the average length of stay and average cost of stay. Specific behaviours can be assessed that are linked to HIS features (e.g., decision support and a change in the completion of screening tests, or more A1c tests ordered for diabetics), as well as specific workflows impacted by HIS implemen-

tation. Intended changes and unintended consequences should be considered when developing an evaluation plan that measures clinical or health behaviour changes, as there may be surprise impacts when workflows are changed.

4.3.4 CAMM Dimension: Clinical (Health) Outcomes

Clinical (Health) Outcomes, the fourth dimension of the CAMM, is defined as impacts that are attributable to the adoption of the HIS. Five aspects of outcomes can be considered when developing measures and metrics for HIS adoption: *patient outcomes*, *provider outcomes*, *organizational outcomes*, *population outcomes* and *cost outcomes*. Outcomes could be considered early or late, depending on evaluation timing. *Patient outcomes* include aspects directly related to individual patient health, such as patient changes in complications due to diabetes. *Provider outcomes* include provider-centric measures, such as better physician retention. *Organizational outcomes* include factors measured at an organizational level (e.g., nosocomial infection rates) whereas *population outcomes* are measured across organizations (e.g., obesity rates, lifespan, myocardial infarction rates). Finally, *cost outcomes* can be considered that describe relative or absolute costs to the healthcare system. The specific outcomes will depend on the HIS, how it is deployed, and the goals of the project. Not all aspects need to be measured.

There can be some confusion or overlap between behavioural changes and early outcomes and there are grey areas between the two. Consider the behaviours as those that are directly related actions under the control of the HIS user. If an electronic medical record recommends that a physician check blood pressures and the rate of blood pressure checking in the office goes up, that is a behaviour change. An early outcome may be a decrease in the values of the blood pressure readings as people then are better managing their blood pressure.

4.4 CAMM Archetypes

CAMM archetypes were developed to help with understanding and applying the CAMM. Archetypes are representational adoption trajectories for health information systems. These would not chart the precise path that an adoption must or would take. Indeed, most real-world adoptions will fall somewhere between two or more of these archetypes. Still, these are helpful illustrations for discussing the ranges of successes, challenges, and failures that can be seen with HIS adoption. The CAMM archetypes are:

1. No Deployment.

2. Low Adoption.

3. Adoption without Benefit (behaviour and outcome).

4. Behaviour Change without Outcome Benefit.

5. Adoption with Outcome Benefits.

6. Benefit without Use.

7. Adoption with Harm.

4.4.1 No Deployment

This archetype describes an HIS initiative that does not reach the end users in a clinical or health setting. No deployment of an HIS can occur for several reasons, including: an incomplete product, lack of funding, strategic change within an organization, significant delays in the product, or unsuccessful testing of a component. Whatever the reason(s), the deployment to end users was stopped prior to a planned go-live event. End users may be involved in the design or testing but there is not a deployment into a real-world setting. This is often the clearest, most obvious archetype.

4.4.2 Low Adoption

In this archetype, the HIS is deployed and available, but availability is followed by minimal or rapidly declining use (Figure 4.2). Users may explore the HIS, but use is not sustained. Without use, it is not reasonable to expect a benefit from the tool. This can been seen with systems that do not support and fit the clinical environment and where use of the HIS is voluntary. This archetype (along with Benefit without Use) highlights the importance of measuring the multiple CAMM dimensions. If only outcomes are measured, one may make an assumption that an intervention is not beneficial even when it is not used.

An example of the Low Adoption archetype would be assessing the impact of decision support alerts in a system that allows users to turn on or off the alerts. An evaluation may show the implementation of specific decision support alerts is not impacting outcomes. If all dimensions were evaluated, it may be found that use was low because most of the users simply turned the alerts off. Without sufficient use, one cannot expect the outcomes to change.

4.4.3 Adoption without Benefit (behaviour and outcome)

Here we see an HIS that is both available and used by end users; however, it is not achieving the intended behaviour changes or the expected outcomes (Figure 4.3). This archetype can be seen when the HIS functions and features do not directly align with the metrics being measured or the HIS features are not sufficiently evidence-based to facilitate the desired behaviour changes and outcomes. It may be seen when the measured clinical behaviours and outcomes are already positive, that there is a ceiling effect, or, conversely, when the clinical environment has limited capacity for change. It can also happen when the timing of the

evaluation is premature and adaptations or changes to health outcomes have not yet occurred.

4.4.4 Behaviour Change without Outcome Benefit

This archetype occurs when an adopted HIS produces the expected changes in behaviour, however the behaviours are not leading to the expected outcomes (Figure 4.4). This can be seen when the intervention isn't sufficiently evidence-based, or the causal chains in the evidence are not sufficient to lead to the outcomes. Again, it may be possible that the outcomes are already good (the ceiling effect) or that the duration or timing of evaluation is too short to see the outcomes. It is important to note that some clinical outcomes are not immediately evident. Successful preventive care programs may not be expected to show benefits in mortality for many years, as the natural history of several diseases are described in years or decades. Thus, early surrogate markers that are connected to evidence are often chosen to support stakeholders in their decision-making.

4.4.5 Adoption with Benefits

This is the archetype that HIS adoption programs expect and hope to see: a clear progression of HIS availability that leads to ongoing use of the HIS (Figure 4.1). HIS use then leads to observable changes in clinical and health behaviours that, in turn, result in improvements in measured outcomes. Note that while the CAMM suggests a causal link between each of the four dimensions, the reality is that healthcare is a complex environment, often with multiple programs seeking the same types of improvements. It is important to note that causation cannot be assumed between the HIS and the outcomes just because they are measured. Some evaluation methods can only describe the correlation of events and the CAMM does not specify evaluation methods. Methods should be sufficiently rigorous to support both the scope of the initiative and ongoing decision-making, in addition to adjustments to the HIS as required over time.

4.4.6 Benefit without Use

Here we see the expected behaviour changes and/or outcomes but *without* the use of the HIS. This occurs (as described above) where there can be multiple overlapping initiatives, each striving to improve the same or similar outcomes (Figure 4.5). Here, another program confounds and impacts the measurements of the HIS behaviour or outcome metrics.

As an example, consider a scenario where a new eHealth tool may be developed to support chronic disease management. Many target users do not use the eHealth tool as they feel it is too cumbersome and their current practices are more efficient. However, at the same time a new funding program for chronic disease management is initiated. This motivates users and many of the chronic disease management activities envisioned to be enabled by the new component are taking place, but through other means. Chronic disease management improves. Clearly, there is a correlation between deployment of the eHealth tool

but there is *not* a correlation with use. This archetype highlights the importance of measuring each of the dimensions as part of an evaluation. Without measuring use, the evaluator and stakeholders could erroneously assume that the HIS is enabling and responsible for the improvement in chronic disease management.

4.4.7 Adoption with Harm

Although we like to focus on benefits, HIS adoption may lead to unintended consequences and harm (Figure 4.6). This archetype highlights the risk of negative effects caused by the use of an HIS. HIS deployments can result in harm from unexpected changes brought about by the implementation of the HIS. Harm can occur from improper design, improper use, or from changes in other workflows (often informal workflows) resulting from the HIS implementation. Potential harm should be considered when planning and should be measured in the evaluation to avoid or limit unintended effects.

4.5 Using the CAMM

The four CAMM dimensions and their aspects help describe trajectories of HIS adoption over time. The CAMM suggests a logical causal chain from availability to use to behaviour changes to resulting changes in outcomes. The CAMM can be helpful in planning evaluations and in explaining findings.

For those who are planning HIS evaluations, the CAMM provides a framework to consider metrics and measures that will change at differing points over time. The CAMM also highlights the need to consider multiple dimensions within an HIS evaluation and when each evaluation dimension will be expected to be most helpful in an adoption's life cycle. Stakeholders will have evaluation needs that have their own timing. The CAMM can help inform and focus the kinds of evaluations that would best support stakeholder needs. It would not be helpful to measure changes in outcomes three months after a diabetes prevention app is published for mobile phones, for example. Those outcomes would not be expected to be measurable for years. Instead, the CAMM would suggest considering metrics for availability (presence on the app stores, presence on smart phones as indicated by number of downloads) and use (number of times the app is opened by how many users, content reviewed). These will show stakeholders meaningful early metrics, which can evolve to the later metrics over time.

The images of the CAMM suggest individual trajectories for each dimension, but the reality is there can often be multiple metrics for each dimension that follow different trajectories. For example, positive and negative outcomes can occur at the same time, depending on the specific metrics an evaluation considers. A targeted intervention may have unintended consequences due to a shifting of resources away from good practice. A particular HIS may have strong areas and weaker areas and thus only measuring its impact in one functional department may fail to present a full picture. Further, adoption of an HIS may be variable

across an organization. The evaluation of an HIS is complex and the CAMM provides an accessible framework to begin planning an evaluation over time.

As an explanatory framework, CAMM can be also applied retrospectively. It can be used to consider the results of an adoption. CAMM can be used with stakeholders to reflect and point to areas of an implementation that should be better explored to understand some results. Availability issues and partial or unexpected use could be discovered in projects where benefits are not being realized. Quantitative and qualitative metrics can be sought retrospectively if needed to help understand an HIS implementation.

4.5.1 Case Study: Using the CAMM to Inform a Personal Health Portal Evaluation

The CAMM was initially developed to help engage stakeholders in the discussion and planning of benefits evaluations for the deployment of a multiphased personal health portal program. This case study will focus on developing an evaluation plan for the initial deployment of the personal health record (PHR) component of a larger Personal Health Portal project. As part of the multi-stakeholder engagement, the goals of the initial PHR deployment were prospectively elicited. The focus of the initial deployment was within a single clinical site and an evaluation plan was developed for this initial deployment, with an eye to stakeholder needs, that included planning for future, broader deployments.

The CAMM dimensions are presented here in "reverse" order as it can be helpful to "start with the end in mind" when developing the evaluation metrics.

4.5.2 Setting: Current State

The site of the initial PHR deployment was a cardiac rehabilitation program for patients who had recently suffered a heart attack. It was a 12-week outpatient program started after patients were stable and discharged from hospital. Patients currently engage with a team of cardiologists, cardiac nurses, dieticians, and exercise therapists to educate and create a personalized program of rehabilitation (which included diet, exercise, medications, monitoring, and self-management) to improve and maintain function. The PHR was being deployed to patients at the start of their program with functions tailored to management of cardiac care and related conditions (e.g., hypertension, diabetes) and a mechanism allowing trusted providers to access the record by virtual check-in.

4.5.3 Predicted Outcomes

From the stakeholder engagement, the key outcome for this deployment was to reduce recurrent heart attacks in patients who had already suffered a heart attack (and thus improve mortality). It was expected that would be achieved through better proximal outcomes like improved blood pressure control, better management of congestive heart failure (CHF), and overall improved patient knowledge of cardiac care and their own care plans.

4.5.4 Expected Behaviours

The stakeholders linked use of the PHR with several health behaviours that would lead to the predicted outcomes. First, the patients would be more engaged, which would result in better blood pressure tracking, improved diet, more exercise, and better adherence to medications. The PHR allowed linking to providers, so stakeholders expected closer and longer follow-up of patients in the cardiac rehabilitation program. This would be seen both in an increase in the number of contacts with the patient and an extension of the rehab programs to more than 12 weeks.

4.5.5 Expected Use

The stakeholders expected patients to use the system regularly to track weight, blood pressure and medication use. They also expected patients to use cardiac rehab self-management plans (e.g., care plans) that were in the PHR. They expected Registered Nurses to log in at least weekly to check on patients in the program. The expectation was that the user experience was easy and intuitive for the patients, thus facilitating self-management.

4.5.6 Expected Availability

The stakeholders assumed availability would be 100% for all participants. Further discussion elicited several specifics: All users (patients and providers) would have access, which included passwords and training; the system would be available through the Internet at points where users expect it to be available (clinic, home); and the PHR had tools available that would support the self-management of patients' cardiac care issues.

4.5.7 Evaluation Metrics and Results

Timing of the deployment and evaluation was relatively short (12 weeks). This was necessary as a key decision was to be made by the steering committee on future deployments within four months of this pilot. As a consequence, outcomes could not be selected, as outcome evaluation would have likely resulted in a null result. Thus, the evaluation focused on early dimensions: *Availability* and *Use*. Data was collected through interviews and focus groups at multiple points in time over the 12-week pilot. Description of adoption would be described for patients and then providers.

Patients: *Availability:* All patients had accounts and training. The PHR was running without issue for the 12 weeks; however, some patients expected the PHR could be accessed through smartphones or tablets and it was (at that time) designed for desktop browsers. Not all patients had computers as they had transitioned to tablets. Content included provincial medication dispensing records and whatever information the patient entered. *Use:* Most patients used the PHR regularly as part of the study. The user experience could have been improved through streamlining the navigation and providing more valuable tools in the PHR that would help patients meet their care plan goals (e.g., reminders).

Patients did not note any behaviour changes, as the PHR for them was primarily a documentation tool.

Providers: *Availability:* There were delays in availability. Specifically, accounts were created for providers but the process for connecting providers to patients was challenging due to timeouts. An asynchronous process, the account linking required multiple steps and with patients not logging into the PHR daily and providers perhaps only working part-time, the window to link provider accounts to patient accounts in the PHR proved difficult. *Use:* Provider use was limited by availability. Virtual connections and monitoring had not begun during the pilot.

The use of the CAMM in this case study intentionally highlights the importance of measuring early dimensions of availability and use in implementations. These can facilitate important improvements in the deployment plans to better achieve adoption and expected benefits. In this case study, the findings were used to inform the next planned deployment and the CAMM was used to frame subsequent deployments in this large, phased program.

4.6 Summary

The CAMM is an adoption model that highlights how evaluation of HIS deployments should change over time. The adoption of health information systems can follow a trajectory of linked activities that are described by the four dimensions: availability, use, behaviour, and outcomes. Each of these dimensions can be used to consider when specific metrics should be measured over time during an ongoing evaluation of an HIS deployment.

The CAMM highlights that evaluations early in the adoption process, such as the case study, should focus on early dimensions of availability and use. Later evaluations should not only focus on the later behaviour and outcomes dimensions, but also should include some assessment of availability and use to ensure that the outcomes are not being seen without the expected adoption of the tools.

4.6.1 Acknowledgements

We would like to thank the Alberta Ministry of Health for supporting the evaluation program for the Patient Health Portal that led to the development of the CAMM.

References

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: A ten-year update. *Journal of Management Information Systems*, *19*(4), 9–30.

Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., & Kyriakidou, O. (2004). Diffusion of innovations in service organizations: Systematic review and recommendations. *The Milbank Quarterly, 82*(4), 581–629.

Hassenzahl, M., & Tractinsky, N. (2006). User experience: A research agenda. *Behavior and Information Technology*, *25*, 91–97.

Holden, R. J., & Karsh, B. T. (2010). The technology acceptance model: Its past and its future in health care. *Journal of Biomedical Infomatics*, *43*, 159–172.

Lee, Y., Kozar, K. A., & Larsen, K. R. T. (2003). The technology acceptance model: Past, present, and future. *Communications of the Association for Information Systems*, *12*, 752–780.

Palacio, C., Harrison, J. P., & Garets, D. (2010). Benchmarking electronic medical records initiatives in the U.S.: A conceptual model. *Journal of Medical Systems*, *34*(3), 273–279.

Pettit, L. (2013). Understanding EMRAM and how it can be used by policy-makers, hospital CIOs and their IT teams. *World Hospitals and Health Services*, *49*, 7–9.

Price, M., & Lau, F. (2014). The clinical adoption meta-model: a temporal meta-model describing the clinical adoption of health information systems. *BMC Medical Informatics and Decision Making, 14*, 43. Retrieved from http://www.biomedical.com/1472-6947/14/43

Price, M., Lau, F., & Lai, J. (2011). Measuring EMR adoption: a framework and case study. *Electronic Healthcare, 10*(1), e25–e30.

Rogers, E. M., & Shoemaker, F. F. (1971). *Communication of innovations: A cross-cultural approach*. New York: The Free Press.

van de Wetering, R., & Batenburg, R. (2009). A PACS maturity model: A systematic meta-analytic review on maturation and evolvability of PACS in the hospital enterprise. *International Journal of Medical Informatics*, *78*(2), 127–140.

van de Wetering, R., Batenburg, R., & Lederman, R. (2010). Evolutionistic or revolutionary paths? A PACS maturity model for strategic situational planning. *International Journal of Computer Assisted Radiology and Surgery, 5*(4), 401–409.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *Management Information Systems Quarterly*, *27*(3), 425–478.

Wu, S., Chaudhry, B., Wang, J., & Maglione, M. (2006). Systematic review: Impact of health information technology on quality, efficiency, and costs of medical care. *Annals of Internal Medicine*, *144*(10), E12–E22.

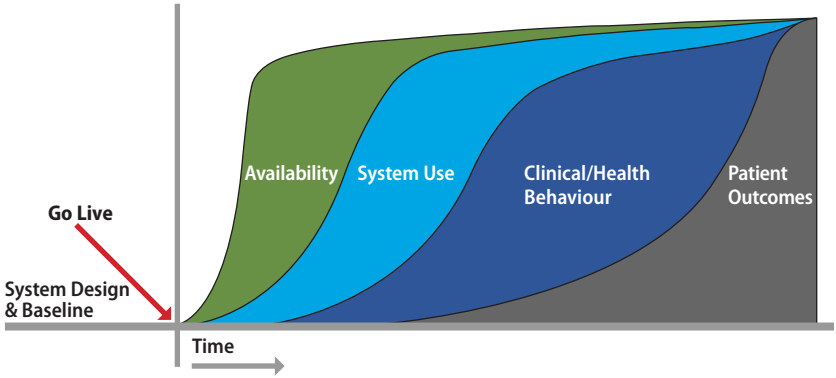


Figure 4.1. The clinical adoption meta-model.

Note. From "The clinical adoption meta-model: a temporal meta-model describing the clinical adoption of health information systems," by M. Price and F. Lau, 2014, *BMC Medical Informatics and Decision Making, 14*(43), p. 2. Copyright 2014 by Price and Lau. CC BY Licence.

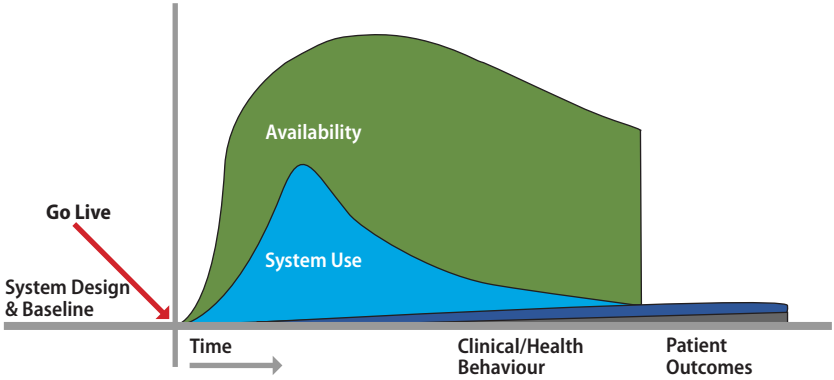


Figure 4.2. Low adoption archetype.

Note. From "The clinical adoption meta-model: a temporal meta-model describing the clinical adoption of health information systems," by M. Price and F. Lau, 2014, *BMC Medical Informatics and Decision Making, 14*(43), p. 5. Copyright 2014 by Price and Lau. CC BY Licence.

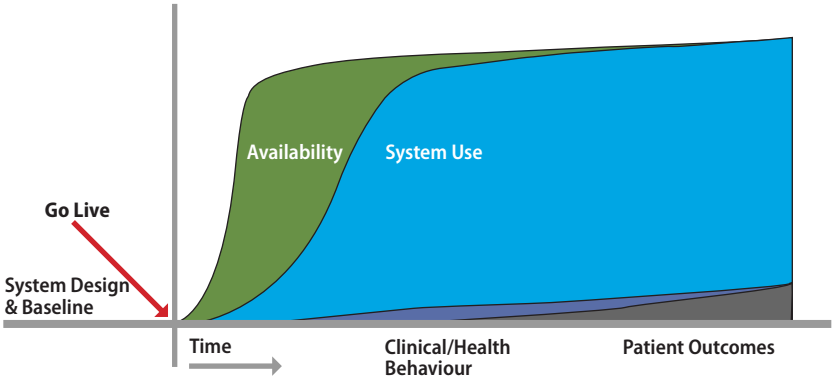


Figure 4.3. Adoption without benefits archetype.

Note. From "The clinical adoption meta-model: a temporal meta-model describing the clinical adoption of health information systems," by M. Price and F. Lau, 2014, *BMC Medical Informatics and Decision Making, 14*(43), p. 5. Copyright 2014 by Price and Lau. CC BY Licence.

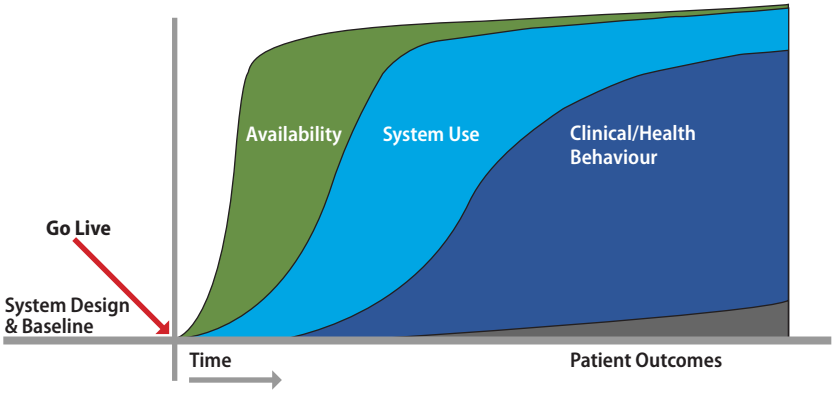


Figure 4.4. Behaviour change without outcome benefits archetype.

Note. From "The clinical adoption meta-model: a temporal meta-model describing the clinical adoption of health information systems," by M. Price and F. Lau, 2014, *BMC Medical Informatics and Decision Making, 14*(43), p. 6. Copyright 2014 by Price and Lau. CC BY Licence.

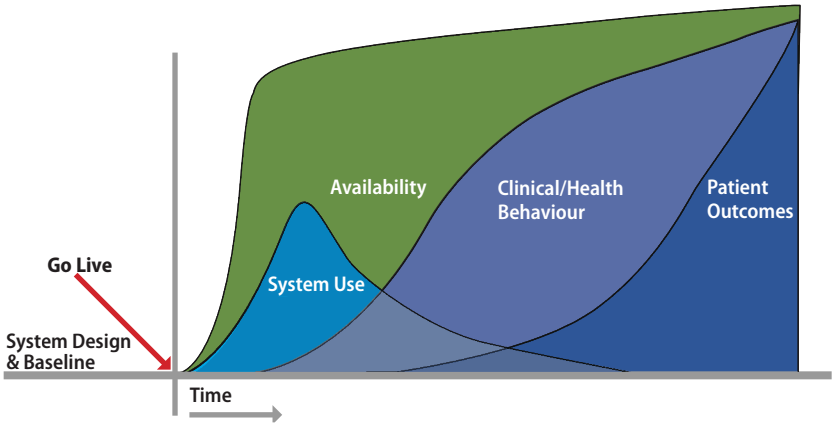


Figure 4.5. Benefit without use archetype.

Note. From "The clinical adoption meta-model: a temporal meta-model describing the clinical adoption of health information systems," by M. Price and F. Lau, 2014, *BMC Medical Informatics and Decision Making, 14*(43), p. 6. Copyright 2014 by Price and Lau. CC BY Licence.

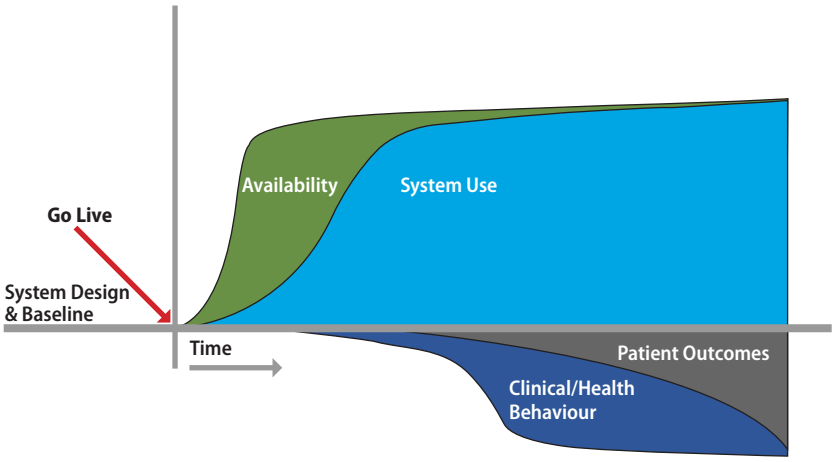


Figure 4.6. Adoption with harm archetype.

Note. From "The clinical adoption meta-model: a temporal meta-model describing the clinical adoption of health information systems," by M. Price and F. Lau, 2014, *BMC Medical Informatics and Decision Making, 14*(43), p. 7. Copyright 2014 by Price and Lau. CC BY Licence.

Chapter 5

eHealth Economic Evaluation Framework

Francis Lau

5.1 Introduction

Increasingly, healthcare organizations are challenged to demonstrate the worth of eHealth investments with respect to their economic return. Over the years, different approaches have been applied to determine the value of eHealth investments such as the financial benefit, cost-effectiveness and quality-adjusted life years gained. Despite the work done to date, there is still limited evidence on the economic return associated with the myriad of eHealth systems deployed. This available evidence is often mixed as to whether eHealth can demonstrate a positive return on the investment or not. Moreover, the methodological rigour of some evaluation studies is questionable.

In 2013, Bassi and Lau published a scoping review of primary studies on the economic evaluation of HIS or health information systems (2013). Based on 33 high-quality HIS economic evaluation studies published between 2000 and 2012 we reported on the key components of an HIS economic evaluation study, the current state of evidence on economic return of HIS, and a set of guidance criteria for conducting HIS economic evaluation studies. Drawing on the review findings, we proposed an economic evaluation classification scheme that is the basis of the eHealth Economic Evaluation Framework described in this chapter.

This chapter describes an eHealth Economic Evaluation Framework based on our scoping review findings and related best practices in economic evaluation literature. The chapter covers the underlying conceptual foundations and the six dimensions of our framework, guidance on its potential use and implications for healthcare organizations.

5.2 Conceptual Foundations

Different economic evaluation approaches for eHealth have been described in the literature. They vary according to the analytical methods applied, the health consequences being considered, and whether it involves a synthesis of multiple studies. The quality of eHealth economic evaluation studies also varies depending on the methodological rigour applied in their design, analysis and reporting. The type and quality of eHealth economic evaluation studies are described below.

5.2.1 Types of Economic Evaluation in eHealth

Economic evaluation is the comparative analysis of alternative interventions with respect to their costs and consequences. Economic evaluation can be based on empirical trials, mathematical models, or a combination of both. The types of economic evaluation studies found in eHealth literature include cost-benefit analysis, cost-effectiveness analysis, and cost-utility analysis. Other variants are cost-minimization analysis, cost-consequence analysis, input cost analysis, and cost-related outcome analysis. These types of economic evaluation are defined below (Roberts, 2006).

- *Cost-benefit analysis* – examines both costs and consequences in monetary terms.
- *Cost-effectiveness analysis* – examines costs and a single consequence in its natural unit such as hospital length of stay in days or frequency of adverse events as a percentage.
- *Cost-utility analysis* – examines costs and a single consequence in the form of a health-related quality of life measure such as quality-adjusted life years.
- *Cost-consequence analysis* – examines the costs and multiple consequences in their natural units without aggregation into a single consequence.
- *Cost-minimization analysis* – examines the least costly consequence among alternatives with equivalent consequences.
- *Input cost analysis* – examines the costs of all alternatives but not their consequences.
- *Cost-related outcome analysis* – examines the consequences of all alternatives in monetary terms but not the input costs incurred.

When the economic analysis involves the comparison of both the costs and consequences, it is considered a full economic evaluation. Cost-benefit, cost-

effectiveness, cost-utility, and cost-consequence analyses are examples of full economic evaluation. If the analysis involves only the costs (e.g., input cost analysis) or consequences (e.g., cost-related outcome analysis), it is considered a partial or one-sided economic evaluation. Cost-minimization is a form of input cost analysis since it assumes all of the consequences are equivalent and therefore the focus is on the least costly alternative.

In the eHealth literature, sometimes the term "benefit" is used to include different types of consequences which may be non-monetary in nature. An example is the term "benefits evaluation" where the benefits can be in dollar terms or in some other units such as hospital length of stay in days or number of adverse events in a given time period. To avoid confusion it is important to describe the type of economic analysis used and the nature of the benefits involved.

5.2.2 Quality of eHealth Economic Evaluation Studies

Different criteria for assessing the quality of economic evaluation studies in terms of their design, analysis and reporting have been published in the literature. In this section, we briefly describe the quality assessment criteria from our scoping review and the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) publication guidelines (Husereau et al., 2013) as two ways to enhance the rigour of our eHealth Economic Evaluation Framework. These are described in more detail in chapter 14 under methodological considerations and best practice guidelines.

Ten quality criteria derived from four literature sources were used in our scoping review to assess the methodological quality of the selected HIS economic evaluation studies (Drummond & Jefferson, 1996; Centre for Reviews and Dissemination [CRD], 2009; Evers, Goossens, de Vet, van Tulder, & Ament, 2005; Machado, Iskedjian, & Einarson, 2006). Each criterion scores between 0 and 1, from not stated, somewhat stated, to clearly stated, for a maximum score of 10 as having the highest quality. These criteria are listed below:

- Is there a research question or definition of the study aim?
- Are the primary outcome measures stated?
- Is the study sample provided and described?
- Is the HIS being evaluated described?
- Is the study time horizon stated?
- Are the data collection methods described?
- Are the analytical methods described?

- Are the results clearly reported with caveats where needed?
- Do the conclusions follow from the study question/objective?
- Are generalizability issues addressed along with limitations?

The CHEERS guidelines were published in 2013 (Husereau et al., 2013) by the Good Reporting Practices Task Force of the International Society for Pharmacoeconomics and Outcomes Research (ISPOR). The guidelines are recommendations for optimized reporting of health economic evaluation studies. They were derived from previous systemic reviews and surveys of task force members, followed by two rounds of Delphi process that reduced an initial list of 44 candidate reporting items to 24 items with accompanying recommendations in a checklist format.

- *Title and abstract* – two items on having a title that identifies the study as an economic evaluation, and a structured summary of objectives, perspective, setting, methods, results and conclusions.
- *Introduction* – one item on study context and objectives, including its policy and practice relevance.
- *Methods* – 14 items on target populations, setting, perspective, comparators, time horizon, discount rate, choice of health outcomes, measurement of effectiveness, measurement and valuation of preference-based outcomes, approaches for estimating resources and costs, currency and conversion, model choice, assumptions, and analytic methods.
- *Results* – four items on study parameters, incremental costs and outcomes, describing uncertainty, and describing heterogeneity.
- *Discussion* – one item on findings, limitations, generalizability and current knowledge.
- *Others* – two items on the source of study funding and conflicts of interest.

5.3 Framework Dimensions

The eHealth Economic Evaluation Framework is derived from our scoping review of HIS economic evaluation studies. Its intent is to provide a classification scheme for the different approaches used in eHealth economic evaluation studies. The framework is made up of six components: having a perspective, options,

time frame, costs, outcomes, and method of analyzing/comparing options. These components are shown in Figure 5.1 and described below.

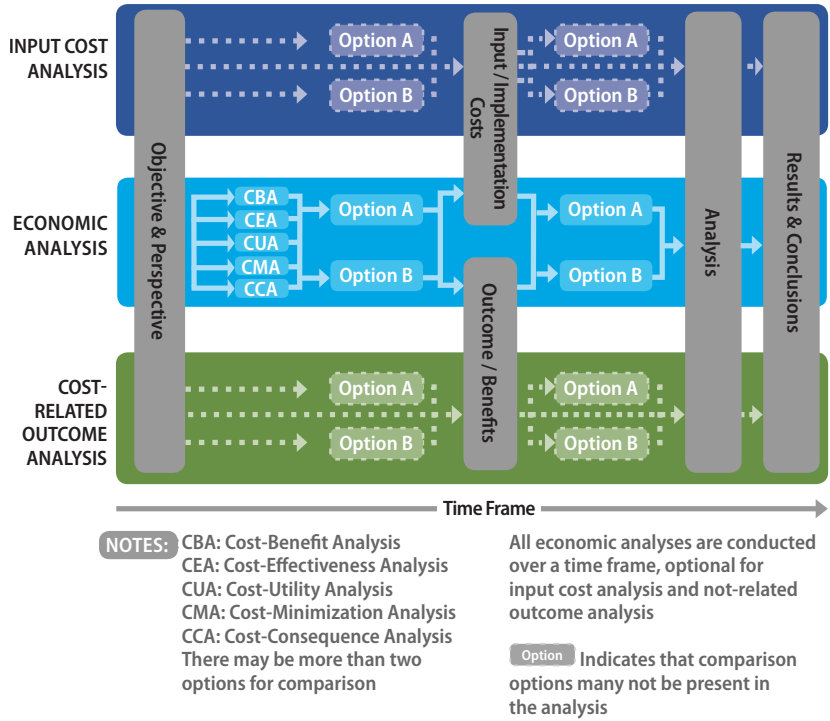


Figure 5.1. eHealth economic evaluation framework.

Note. From "Measuring value for money: A scoping review on economic evaluation of health information systems," by J. Bassi and F. Lau, 2013, *Journal of American Medical Informatics Association, 20*(4), p. 793. Copyright 2013 by Oxford University Press, on behalf of American Medical Informatics Association. Reprinted with permission.

5.3.1 Perspective

Perspective refers to the point of view under which an eHealth economic evaluation is being conducted. It is an important component in the framework because the costs and consequences accrued can affect different parts of the healthcare system. As such, the economic return of an eHealth system is dependent upon who makes the decision, who incurs the costs and who benefits from the consequences. For instance, care providers mandated by the government to adopt an electronic prescription tracking system may perceive it as an added cost that only benefits the government by controlling their practice. The perspectives considered in our framework are those of the individual, organization, payer, and society at large. These are defined below:

- *Individual* – the person affected such as the provider, patient or caregiver. The effect may involve a change in the person's expenditures, routines and/or health conditions.

- *Organization* – the group affected such as the health region, professional association, or patient advocacy where multiple individuals within the group are affected in similar ways.

- *Payer* – the group that finances the healthcare service such as the government or private insurers. The effect may involve a change in the group's cost of providing the service.

- *Society* – the general public affected such as the residents in a geographic region or the entire population of a country. The effect may involve a change to the overall financing of the healthcare system and/or the health status of the population.

5.3.2 Options

Options are the alternative eHealth systems being considered. It is important to clearly define each eHealth system option since they often perform multiple functions and can be adopted for different reasons by different organizations. Also, the behaviour of the system can evolve over time as users become more experienced in using it to support their work. Increasingly, eHealth systems are combined with other interventions to enhance the intended effects. For these reasons, the features within each of these options must be clearly defined for meaningful comparisons to be made. The types of options reported in the eHealth economic evaluation literature are with or without the system, pre- or post-implementation, types of systems, levels of systems, different time points, and different sites. These options are defined below:

- *With or without the system* – one or more eHealth system options and a status quo with no system

- *Pre- or post-implementation* – before and after the adoption of an eHealth system

- *Types of systems* –different eHealth system options with the same or similar functions

- *Levels of systems* – extent of eHealth systems and/or functions adopted in the organization

- *Different time points* – the same eHealth system at different points in time

- *Different sites* – the same eHealth system adopted in different organizations or locations

Two important aspects of options are the status quo and opportunity cost. Status quo refers to the costs and consequences of the current situation without adopting any eHealth systems, or a default "do nothing" position. Opportunity cost refers to the foregone benefit as a result of selecting a given eHealth system option. Status quo and opportunity cost are important in eHealth investment decisions when there are limited resources among competing priorities. For example, a healthcare organization addressing patient medication safety has to decide whether its existing rate of medication errors, or the status quo, is acceptable or needs improvement with an electronic surveillance system as an option. Similarly, an organization wishing to adopt an EMR system to improve its overall care delivery may consider the opportunity cost by asking whether the EMR investment can be better spent elsewhere with comparable effect.

5.3.3 Time Frame

Time frame refers to the length of time for which the costs and consequences of an eHealth system are accrued. One must allow for sufficient time to ensure all of the relevant costs are captured and the consequences are realized as they can accrue differently over time. Often there is a time lag before the consequences, such as a reduction in the rate of adverse events, can be realized after the adoption of an eHealth system. For pragmatic reasons, studies based on empirical data for costs and consequences tend to use shorter time frames, as it is difficult and costly to collect data for a long period. Studies based on mathematical modelling tend to have longer time frames since there is little added effort to predict long-term trends. The time frames reported in eHealth economic evaluation literature are less than one year, one to five years, six to 10 years, and greater than 10 years. They are defined below:

- *Less than one year* – typically for small-scale studies where empirical data on costs and consequences from an eHealth system or intervention are collected over a short time period such as three to six months for comparison. Sometimes the cost and/or consequence are extrapolated to an annual period such as estimated cost savings from an EMR system per year.

- *One to five years* – the most common time periods used are between one and five years in duration to capture the costs, consequences, or both, that are accrued. Sometimes different time periods are used to collect the accrued costs and consequences. For instance one may extract the historical costs for EMR adoption over one year then estimate the return on investment over a five-year period.

- *Six to 10 years* – typically for modelling studies where the costs and consequences are projected over a six- to 10-year period. The data can be based on historical, prospective, estimated or combined sources.

- *Greater than 10 years* – mostly for predicting the long-term consequences of an eHealth system such as the cumulative economic impact of a diabetes management system expected over a 40-year period.

- *Multiple time points* – typically in studies where different types of costs and consequences are captured across multiple time periods depending on the availability of the data.

Note that the time period covered in an economic evaluation study is different from the time it takes to conduct the study itself. For instance, the economic return of a computerized provider order entry system (CPOE) may be determined over a 5-year period to ensure all of the costs incurred are captured and the CPOE is sufficiently stabilized to realize an improvement in ordering medications. Yet the study itself may only take two or three months to collect and analyze the data if it is retrospective in nature or if predictive modelling is used to estimate the effect over a five-year period based on historical data.

5.3.4 Input Costs

Input costs are the amounts of money spent in the adoption of an eHealth system. The types of costs reported in the eHealth economic evaluation literature are one-time direct costs, ongoing direct costs, and ongoing indirect costs. They are defined below.

- *One-time direct costs* – expenditures incurred in order to implement the system. They include such items as hardware equipment, software licences, application development/customization, data conversion, system configuration, training, user and technical support.

- *Ongoing direct costs* – recurrent expenditures to operate the system after its implementation. They include such items as hardware and software maintenance, system upgrades, technical and support staffing, ongoing training, and related professional services (e.g., system audits).

- *Ongoing indirect costs* – recurrent expenditures related to the system that is allocated by the organization after its implementation. They

include prorated expenditures such as managing IT-related privacy, security, policy and help desk, and changes in staff workload.

Intangible costs are another type of cost that is mentioned in the economic evaluation literature. Intangible costs refer to things that are unquantifiable or difficult to measure. Examples of intangible costs in the adoption of an eHealth system are a change in staff morale and patient anxiety before, during and after the implementation of an EMR as they learn to work with the new system. Intangible costs are seldom addressed in eHealth economic evaluations. One approach is to estimate intangible costs as a type of input or outcome such as the quality of staff work life in terms of productivity before or after the adoption of an EMR.

5.3.5 Outcomes

Outcomes refer to the consequences from adopting an eHealth system. There are different types of outcomes reported in the eHealth economic evaluation literature. These outcomes may be financial or non-financial in nature, and can be derived from empirical data, projections or both. Financial outcomes include changes in revenues, labour and supply costs, and capital costs expressed in monetary units. Non-financial outcomes include changes in resource utilization and health outcomes in their natural units. These types of consequences are outlined below. Note that only tangible outcomes are considered here.

- *Revenues* – money generated from billing and payment of patient care service provision supported by the eHealth system, and change in such financial arrangements as the reimbursement rates, accounts receivable days and payer mix for service claims.

- *Labour and supply cost savings* – change in staffing costs due to altered productivity associated with the eHealth system such as data entry, charting, communication and reporting, and change in supply costs such as the amount of stocked materials and goods consumed.

- *Capital cost savings* – change in capital expenditures for such items as facilities, equipment and technology due to the adoption of an eHealth system.

- *Resource utilization* – change in healthcare resource usage such as the volume of laboratory and radiology tests, medications and other diagnostic/interventional procedures consumed.

- *Health outcomes* – change in patients' conditions and clinical events such as one's physiologic status, or the number of medical errors and adverse events reported.

5.3.6 Comparison of Options

Comparison of options refers to the analytical methods used to determine the return on investment for each eHealth system option being considered. Different methods have been reported in the eHealth economic evaluation literature. They include accounting, statistical, and operations research methods that draw on different types of data as their input sources. These are defined below.

- *Data sources* – tabulation of cost and outcome data as the input data sources, based on historical records, expert estimates, model projections, or combinations.

- *Accounting* – measuring the financial performance of each option, which includes the outcome measures, time value of money, uncertainty and risks. Examples of outcome measures are incremental cost-effectiveness ratio, payback, net present value, operating margin and quality-adjusted life years. Examples of time value of money are discounting, inflation, depreciation and amortization. Examples of handling uncertainty and risks are sensitivity and scenarios analysis.

- *Statistics* – measuring the financial performance of each option based on statistical techniques such as linear/logistic regression, general linear modelling and testing for group differences.

- *Operations research* – measuring the financial performance of each option based on operational research methods such as panel regression, parametric cost analysis, stochastic frontier analysis and simulation modelling.

5.4 Framework Usage

The eHealth Economic Evaluation Framework was derived from a scoping review of 33 high-quality HIS economic evaluation studies published between 2000 and 2012. The review provides a rich source of published studies, methods, measures, and lessons that can serve as guidance for designing, analyzing and reporting eHealth economic evaluation studies. The potential usage and implications of this framework based on the six components reported in the review are described below.

5.4.1 Potential Usage

For the 33 HIS studies in the review, 12 were considered full economic evaluations as they included all six framework components. Of these, six were on cost-benefit, two were on cost-effectiveness, two on cost-consequence, and one on cost-utility. For the remaining studies, 16 were on cost-related outcomes, and five on input costs. As for the categories described under each of the six framework components, their patterns of usage among the 33 HIS studies are tabulated below (for more detail, see Bassi & Lau, 2013).

- *Perspective* – Most studies (87.9% or 29/33) were based on an organizational perspective, while 12.1% (4/33) were on society, and 3.0% (1/33) each on individual and payer. Note that the total count exceeds 100% as two studies had two perspectives each and therefore were counted twice.

- *Time Frame* – Over half (54.5% or 18/33) of the studies had time periods of one to five years. Of the remainder, 24.2% (8/33) had six to 10 years, 12.1% (4/33) less than one year, and 3.0% (1/33) each for less than six months, greater than 10 years and multiple time points, respectively.

- *Options* – Close to half (45.5% or 15/33) of the studies had options of with or without the system, while 27.3% (9/33) had pre- and post-implementation options. The remaining were 9.1% (3/33) on different types of options with similar functions, 9.1% for different levels of adoption, and 3.0% (1/33) each for different time points and not defined, respectively.

- *Input Costs* – 277 measures were reported based mostly on input cost analysis and cost benefit analysis studies. The majority of these measures were one-time direct costs (60.6% or 168/277) with the remaining as ongoing direct costs (32.9% or 91/277). Ongoing indirect costs were seldom mentioned (0.7% or 2/277). Of the 168 one-time direct cost measures, just over one-third (35.1% or 59/168) were for application development and deployment, with the remaining on hardware and software (32.1% or 54/168), initial data collection/conversion (6.5% or 11/168), initial user training (6.0% or 10/168) and other costs. Of the 91 ongoing direct cost measures, close to half were for IT and support staff salaries (25.3% or 23/91) and software licences, maintenance and upgrades (20.0% or 18/91). Many studies also had direct and indirect costs combined into other, overall and total costs.

- *Outcomes* – 195 measures were reported based mostly on cost-benefit analysis or cost-related outcome analysis studies. Close to half of the measures (46.7% or 91/195) were on resource utilization, mostly for medications (47.3% or 43/91) and laboratory tests (34.1% or 31/91). Other outcome categories include labour savings (17.4% or 34/195), healthcare service provision savings (12.3% or 24/195), and total costs/savings (12.8% or 25/195). Examples of labour savings reported are efficiency and time-related savings. Healthcare service provision savings refer to changes in clinical outcomes and include rates of adverse drug events, patient safety events and disease prevention or management. Total cost savings include such measures as annual cost savings, net benefit and incremental cost effectiveness ratio.

- *Comparison of Options* – Accounting was the most common method (72.7% or 24/33) used to compare options through such measures as the incremental cost-effectiveness ratios, return on investment, payback, net present value, net benefit, operating margin, least cost, average cost and cost savings. The estimation methods used to estimate future outcomes included linear/logistic regression (15.2% or 5/33), scenarios analysis (9.1% or 3/33), and general/linear modelling (9.1%). Many studies adjusted for inflation (30.3% or 10/33), discounting (24.2% or 8/33), and amortization/depreciation (12.1% or 4/33). Some studies applied statistical methods to test for differences among groups such as *t*-test (15.2% or 5/33), analysis of variance (6.1% or 2/33) and chi-square (3.0% or 1/33). Several studies used econometric or financial modelling methods based on simulation (12.1% or 4/33), parametric cost analysis (6.1% or 2/33), stochastic frontier analysis (3.0% or 1/33), and panel regression (3.0% or 1/33). For data sources, close to half (48.5% or 16/33) of the studies used both historical and published costs for comparison, while just over one-tenth (12.1% or 4/33) used historical and estimated costs. The remaining studies (39.4% or 13/33) used historical and estimated costs to project future costs and benefits.

5.5 Implications

The eHealth Economic Evaluation Framework described in this chapter can serve as a classification scheme for the approaches used to evaluate the economic return on eHealth system investments. The framework defines the six key components that should be addressed when designing, analyzing and reporting eHealth economic evaluation studies. There are four practice implications to be considered when applying this framework: (a) the type of economic

analysis involved, (b) the use of estimated costs, (c) the importance of incremental return, and (d) the issue of opportunity cost (Gospodarevskaya & Westbrook, 2014). These are described below.

- *Type of economic analysis* – From the scoping review we found only 12 studies were considered full economic evaluations with half of them being cost-benefit analysis, while the other types such as cost-utility analysis were rarely seen. The review included 16 cost-related outcome analysis studies that focused mostly on cost savings or cost changes after implementation. However, without knowing the initial costs of implementing the system it is difficult to determine whether the savings were worth the investment. Similarly, there were five studies on input cost analysis, which alone does not reveal the respective return on each option to make an investment decision.

- *Estimated costs* – Over half of the 33 studies in the review included some type of estimated costs when deriving the input costs or projecting cost-related outcomes. In general, economic evaluation studies that are based on expert opinions, cost avoidance and modelling should be viewed with caution. Expert opinions are subjective in nature and it is often difficult to validate their accuracy. Cost avoidance refers to potential reductions only and these are less convincing than tangible measureable outputs such as actual cost savings in dollars. Modelling studies are hypothetical in nature and may lead to unrealistic forecasts and expectations.

- *Marginal return* – Within the economics discipline, full economic evaluation is the comparative analysis of options that involves the identification, measurement and valuation of costs and outcomes to determine the incremental difference in costs in relation to difference in outcomes. This is demonstrated through the cost for each additional unit of outcome compared with an incremental cost-effectiveness ratio (ICER). In the review, only two studies applied ICER to determine the incremental return of the investment decision. The remaining studies compared the costs and outcomes for each option, which provides only an average cost-effectiveness ratio.

- *Opportunity cost* – Related to the notion of incremental return is the opportunity cost, which is the foregone benefit from the alternative use of resources beyond the eHealth system options. As such, the investment decision must demonstrate its economic efficiency by providing better value than the alternative use of resources and associated outcomes, including non-eHealth options.

5.6 Summary

This chapter described the eHealth Economic Evaluation Framework as a classification scheme to help understand the different approaches used in eHealth economic evaluation studies. The framework has six components: having a perspective, options, time frame, input costs, outcomes, and method of analyzing/comparing options. Best practice guidance does exist for each of the six framework components and there are quality criteria for assessing such studies that should be considered. By applying the framework components one can improve the design, analysis, and reporting of eHealth economic evaluation.

References

Bassi, J., & Lau, F. (2013). Measuring value for money: A scoping review on economic evaluation of health information systems. *Journal of American Medical Informatics Association, 20*(4), 792–801.

Centre for Reviews and Dissemination (CRD). (2009, January). *Systematic reviews — CRD's guidance for undertaking reviews in health care.* York, UK: University of York. Retrieved from http://www.york.ac.uk/crd/guidance

Drummond, M. F., & Jefferson, T. O. (1996). Guidelines for authors and peer reviewers of economic submissions to the BMJ. *British Medical Journal, 313*(7052), 275–283.

Evers, S., Goossens, M., de Vet, H., van Tulder, M., & Ament, A. (2005). Criteria list for assessment of methodological quality of economic evaluations: Consensus on health economic criteria. *International Journal of Technology Assessment in Health Care*, *21*(2), 240–250.

Gospodarevskaya, E., & Westbrook, J. I. (2014). Correspondence: Call for discussion about the framework for categorizing economic evaluations of health information systems and assessing their quality. *Journal of American Medical Informatics Association, 21*(1), 190–191.

Husereau, D., Drummond, M., Petrou, S., Carswell, C., Moher, D., Greenberg, D., … Loder, E. (2013). Consolidated health economic evaluation reporting standards (CHEERS) – Explanation and elaboration: A report of the ISPOR health economic evaluation publication guidelines good reporting practices task force. *Value in Health, 16*, 231–250.

Machado, M., Iskedjian, M., & Einarson, T. R. (2006). Quality assessment of published health economic analyses from South America. *Annals of Pharmacotherapy, 40*(5), 943–949.

Roberts, M. S. (2006). Economic aspects of evaluation. In C. P. Friedman & J. C. Wyatt (Eds.), *Evaluation methods in biomedical informatics* (2nd ed., pp. 301–337). New York: Springer.

Chapter 6

Pragmatic Health Information Technology Evaluation Framework

Jim Warren, Yulong Gu

6.1 Introduction

This chapter outlines a pragmatic approach to evaluation, using both qualitative and quantitative data. It emphasizes capture of a broad range of stakeholder perspectives and multidimensional evaluation on criteria related to process and culture, as well as outcome and IT system integrity. It also recommends underpinning quantitative analysis with the transactional data from the health IT systems themselves. The recommended approach is iterative and Action Research (AR) oriented. Evaluation should be integral to implementation — it should begin, if possible, before the new technology is introduced into the health workflow and be planned for along with the planning of the implementation itself. Evaluation findings should be used to help refine the implementation and to evoke further user feedback. Dissemination of the findings is also integral and should reach all stakeholders considering uptake of similar technology.

This Health Information Technology (IT) Evaluation Framework was developed under the commission of the New Zealand (N.Z.) National Health IT Board to support implementation of the New Zealand National Health IT Plan (IT Health Board, 2010) and health innovation in the country in general. The framework was published in 2011 by the N.Z. Ministry of Health (Warren, Pollock, White, & Day, 2011) with a summary version of this report presented in the Health Informatics New Zealand 10th Annual Conference and Exhibition (Warren, Pollock, White, Day, & Gu, 2011). This framework provides guidelines intended to promote consistency and quality in the process of health IT evalu-

ation, in its reporting and in the broad dissemination of the findings. In the next section, we discuss key elements of the conceptual foundations of the framework. In the third section we specifically address formulation of a Benefits Evaluation Framework from a broad Criteria Pool. We conclude with the implications of applying such a framework and summary.

6.2 Conceptual Foundations and General Approach

A number of sources informed this framework's recommendations for how to design an evaluation. In a nutshell, the philosophy is:

- Evaluate many dimensions – don't look at just one or two measures, and include qualitative data; we want to hear the "voices" of those impacted by the system.

- Be adaptive as the data comes in – don't let the study protocol lock you into ignoring what's really going on; this dictates an iterative design where you reflect on collected data before all data collection is completed.

6.2.1 Multiple dimensions

Of particular inspiration toward our recommendation to evaluate many dimensions is the work of Westbrook et al. (2007) who took a multi-method sociotechnical approach to health information systems evaluation encompassing the dimensions of work and communication patterns, organizational culture, and safety and quality. They demonstrate building evaluation out of a package of multiple relatively small study protocols, as compared to a central focus on randomized controlled trials (RCTs), as the best source of evidence. Further, a "review of reviews" of health information systems (HIS) studies (Lau, Kuziemsky, Price, & Gardner, 2010) offers a broad pool of HIS benefits which the authors base on the Canada Health Infoway Benefits Evaluation (BE) Framework (Lau, Hagens, & Muttitt, 2007), itself based on the Information Systems Success model (Delone & McLean, 2003). Lau et al. (2010) further expand the Infoway BE model based on measures emerging in their review which didn't fit the existing categories. In addition, our approach is influenced by Greenhalgh and Russell's (2010) recommendation to supplement the traditional positivist perspective with a critical-interpretive one to achieve a robust evaluation of complex eHealth systems that captures the range of stakeholder views.

6.2.2 Grounded Theory (GT) and the Interpretivist view

In contrast to measurement approaches aimed at predefined objectives, GT is an inductive methodology to generate theories through a rigorous research process leading to the emergence of conceptual categories. These conceptual categories are related to each other, and mapping such relationships constitutes a

theoretical explanation of the actions emerging from the main concerns of the stakeholders (Glaser & Strauss, 1967). Perhaps the most relevant message to take from GT is the idea of a "theory" emerging from analysis and acceptance of the messages in the interview data; this contrasts with coming in with a hypothesis that the data tests.

We recommend that the evaluation team (the interviewer, and also the data analyst) allow their perspective to shift between a positivist view (that the interview is an instrument to objectively measure the reality of the situation) and an interpretivist view. Interpretivism refers to the "systematic analysis of socially meaningful action through the direct detailed observation of people in natural settings in order to arrive at understandings and interpretations of how people create and maintain their social worlds" (Neuman, 2003, p. 77). An interpretivist accepts that their presence affects the social reality. Equally importantly, the interpretivist accepts individual views as a kind of reality in their own right. The complex demands, values and interrelationships in the healthcare environment make it entirely possible for different individuals to interpret and react to the exact same health IT system in very different ways. The interpretivist takes the view that each stakeholder's perspective is equally (and potentially simultaneously) valid; the aim is to develop the understanding of why the ostensibly contradictory views are held.

Ideally, in developing themes (or GT conceptual categories) from interview data, one would conduct a complete "coding" of the interview transcripts, assigning each and every utterance to a place in the coding scheme and then allowing a theory to emerge that relates the categories. Further, since this process is obviously subjective, one should regard the emerging schema with "suspicion" and contest its validity by "triangulation" to other sources (Klein & Myers, 1999), including the international research literature. These techniques are illustrated in the context of stakeholders of genetic information management in a study by Gu, Warren, and Day (2011). Creating a complete coding from transcripts is unlikely to be practical in the context of most health IT evaluation projects. As such, themes may be developed from interview notes, directly organizing the key points emerging from each interview to form the categories to subsequently organize into a theory of the impact of the health IT system on the stakeholders. Such a theory can then be presented by describing in detail each of several relevant themes.

6.2.3 Evaluation as Action Research (AR)

An ideal eHealth evaluation has the evaluation plan integrated with the implementation plan, rather than as a separate post-implementation project. When this is the case, the principles of AR (McNiff & Whitehead, 2002; Stringer, 1999) should be applied to a greater or lesser degree. The AR philosophy can be integrated into interviews, focus groups and forums in several ways that recognize that the AR research aims to get the best outcome (while still being a faithful reporter of the situation) and will proceed iteratively in cycles of planning, re-

flection and action. With the AR paradigm — in which case the evaluation is probably concurrent with the implementation — the activities of evaluation can be unabashedly and directly integrated with efforts to improve the effectiveness of the system.

With respect to AR, at the minimum allow stakeholders, particularly end users of the software, to be aware of the evaluation results on an ongoing basis so that they: (a) are encouraged by the benefits observed so far, and (b) explicitly react to the findings so far to provide their interpretation and feedback. At the most aggressive level, one may view the entire implementation and concurrent evaluation as an undertaking of the stakeholders themselves, with IT and evaluation staff purely as the facilitators of the change. For instance, Participatory Action Research (PAR) methodology has been endorsed and promoted internationally as the appropriate format for primary health care research and, in particular, in communities with high needs (Macaulay et al., 1999). PAR is "based on reflection, data collection, and action that aims to improve health and reduce health inequities through involving the people who, in turn, take actions to improve their own health" (Baum, MacDougall, & Smith, 2006, p. 854). This suggests an extreme view where the *patients* are active in the implementation; a less extreme view would see just the healthcare professionals as the participants.

Even when the evaluation is clearly following the formal end of implementation activities (which, again, is not ideal but is often the reality), an AR philosophy can still be applied. This can take the form of the evaluation team:

- Seeking to share the findings with the stakeholders in the current system implementation and taking the feedback as a further iteration of the research;
- Actively looking for solutions to problems identified (e.g., adapting interview protocols to ask interviewees if they have ideas for solutions);
- Recommending refinements to the current system in the most specific terms that are supported by the findings (with the intent of instigating pursuit of these refinements by stakeholders).

It is likely that many of the areas for refinement will relate to software usability. It is appropriate to recognize that implementation is never really over (locally or nationally), and that software is — by its nature — amenable to modification. This fits the philosophy of Interaction Design (Cooper, Reinmann, & Cronin, 2007) which is the dominant paradigm for development of highly usable human-computer interfaces and most notably adhered to by Apple Incorporated. Fundamental to Interaction Design is the continuous involvement of users to shape the product, and the willingness to shape the product in response to user feedback irrespective of the preconceptions of others (e.g., man-

agement and programmers). If possible, especially where evaluation is well integrated with implementation, Interaction Design elements should be brought to bear as part of the AR approach.

A corollary to recommending an AR approach as per above is that the evaluation process is most appropriately planned and justified along with the IT implementation itself. This leads to setting aside the appropriate resources for evaluation, and creates the expectation that this additional activity stream is integral to the overall implementation effort.

6.3 Benefits Evaluation Framework

There is a wide range of potential areas of benefit (or harm) for IT systems in health, constituting a spectrum of targets for quantitative and qualitative assessment. The specific criteria for a given evaluation study should not be chosen at random. Rather, the case for what to measure and report should be carefully justified. There are several types of sources that can inform the formulation of a benefits framework for a given evaluation study:

- Necessary properties – for systems within the scope of this framework, which touch directly on delivery of patient care, it is difficult to see how patient safety can be omitted from consideration. Also health workforce issues, such as user satisfaction with the system, are difficult to ignore (at least in terms of looking out for gross negative effects).

- Standards and policies – the presence of specific functions or achievement of specific performance levels may be dictated by relevant standards or policies (or even law).

- Academic literature and reports – previous evaluations, overseas or locally, may provide specific expectations about benefits (or drawbacks to look out for).

- Project business case – most IT-enabled innovations will have started with a "project" tied to the implementation of the IT infrastructure, or a significant upgrade in its features or extension in its use. This project will frequently include a business case that promises benefits that outweigh costs, possibly with the mapping of benefits into a financial case. The evaluation should assess the key assertions and assumptions of the business case.

- Emergent benefits – ideally the evaluation should be organized with an iterative framework that allows follow-up on leads; for example, initial interviews might indicate user beliefs about key benefits of

the system that were outside the initial benefits framework and which could then be confirmed and measured in quantitative data.

With respect to the last point above, the benefits framework may evolve over the course of evaluation, particularly if the evaluation involves multiple sites or spans multiple phases of implementation. Thus, the benefits framework may start with the business case assumptions and a few key standards and policy requirements, plus necessary attributes about patient safety and provider satisfaction; it may then evolve after initial study to include benefits that were not explicitly anticipated prior to the commencement of evaluation.

From the sources cited in 6.2.1 above, and our own experience, we draw the criteria pool in Table 6.1. Evaluators should select a mix of criteria from the major dimensions of this pool in identifying evaluation measures for a specific evaluation project. The major focus should be on criteria from the *Impact* genre. Areas that cannot be addressed in depth (which will almost always be most of them) should be addressed qualitatively within the scope of stakeholder interviews. Some areas, such as direct clinical outcomes, are likely to be beyond the scope of most evaluation studies. Moreover, criteria from the criteria pool may be supplemented with specific functional and non-functional requirements that have been accepted as critical success factors for the particular technology in question.

Table 6.1
Criteria Pool

Criteria Domain	Criteria Type	Examples / Comments
Genre: Impact		
Work and Communication Patterns	Efficiency	Time-and-motion measurements, logging of screen access times, transactional log cycle times (e.g., received-to-actioned latency), direct expenditure (staff time or materials), self-report of task time, impression of efficiency; also Safety and Quality or Clinical Effectiveness (see below) of a given resource
	Coherence	Interruptions, multi-tasking (observed or self-reported)
Organizational Culture	Positivity	Reporting feeling positive / motivated, sick leave rates, turnover
	Safety (culture of)	Reported feeling that system is safe, specific safety promoting practices (e.g., incident reporting and review) – also see Safety and Quality domain below
	Effectiveness and Quality (culture of)	Self-report that efforts are effective / that quality matters, quality improvement activity
	Social networks	Levels of inter-professional communication, inter-professional trust, respect and empathy
	Patient centredness	Patient engagement, adherence, confidence, knowledge
Safety and Quality	Safety	Incident rates, timeliness of review, potential sources of error including data inaccuracy (wrong patient details, incorrect / missing / duplicate clinical data) and illegibility; also see Clinical Effectiveness below
	Quality	See Organizational Culture above and Clinical Effectiveness below
Clinical Effectiveness	Outcome	Mortality, morbidity, readmission, length of stay, patient functional status or quality of health/life (e.g., via the 36-item Short Form Health Survey, SF-36)
	Indicator	Glycated haemoglobin (HbA1c), blood pressure, etc.
	Process measure	Clinical practice guideline adherence – also domains above

Table 6.1
Criteria Pool

Criteria Domain	Criteria Type	Examples / Comments
Genre: Product		
IT System Integrity	Stability	Uptime, errors (logged or self-report), disaster recovery features, maintenance effort
	Data quality	See Safety above
	Data security	IT expert opinion, standards compliance, evidence of breaches
	Standards compliance	International / national compliance, demonstrated interoperability
	Scalability	Response time, maintainability / tailorability / extensibility, IT expert opinion
Usability	Uptake / Use	Rate and extent of uptake, persistence of use of alternatives / workarounds (as measured from transactional systems, or self-report)
	Efficiency	As per Impact genre above
	Accuracy	Data entry / interpretation error rates – as per Safety above
	Learnability	Extent of feature use, help desk requests, rate of uptake
	Satisfaction	Overall happiness with solution (e.g., desire to continue using it)
Vendor Factors		Cost competitiveness of licensing / services, vendor support / commitment

Table 6.1
Criteria Pool

Criteria Domain	Criteria Type	Examples / Comments
Genre: Process		
Project Management		On time, on budget, with proposed features / benefits
Participant Experience		Disruption (self-report or using intermediate measures from the Impact genre), angst / anger, meeting expectations, feeling included, impact on Organizational Culture (as per Impact genre)
Leadership and Governance		Identification of leaders, ability to have bridged difficult transitions, role in maintaining quality of Participant Experience and meeting Project Management goals

Note. From *A framework for health IT evaluation*, by J. Warren, M. Pollock, S. Day, and Y. Gu, 2011, pp. 3–4. Paper presented at the Health Informatics New Zealand 10th Annual Conference and Exhibition, Auckland. Copyright 2011 by Health Informatics Conference. Reprinted with permission.

6.4 Guidance on Use of the Framework and its Implications

6.4.1 Guidance

To achieve a multidimensional evaluation, and best leverage the available data sources, it is recommended that an evaluation of health IT implementation include at least the following elements in the study's data collection activities:

- Analysis of documents, physical system and workflow.
- Semi-structured interviewing and thematic analysis of interview content. This may take the form of one-on-one interviews or focus groups, or (ideally) a combination, and should take an iterative, reflective and interpretivist approach.
- Analysis of transactional data, that is, analysis of the records that result from the direct use of information systems in the implementation setting(s).

The findings from these data sources will support assessment with respect to criteria selected from the criteria pool listed in Table 6.1.

Two further elements of study design are essential:

- Assessment of patient safety – at least insofar as to ask stakeholders working at the point of care to explain how the implementation may be improving or threatening safety.

- Benefits framework – to collect data that supports a defensibly appropriate assessment, the performance expectations should be defined working from criteria as per section 6.3 above; in keeping the GT and AR, these criteria (and thus the focus of evaluation) may be allowed to adjust over the course of the project (obviously with the agreement of funders each step of the way!).

Evaluation may also involve questionnaires and timed observations (automatically or manually). Defining a control group is optional but valuable to make a more persuasive case with respect to the innovative use of IT indeed being the source of quantitative changes in system performance. A pragmatic level of control may be to draw parallel data from a health delivery unit with characteristics similar to the one involved in the implementation. It is essential to be clear about what is being evaluated, but it is also essential to match the study design and evaluation objectives to the available resources.

The framework has been tested in the context of evaluations of several regional electronic referral (eReferral) projects in New Zealand. The eReferral reports (Day, Gu, Warren, White, & Pollock, 2011; Gu, Day, Humphrey, Warren, & Pollock, 2012; Warren, Gu, Day, Pollock, & White, 2012; Warren, Pollock, White, & Day, 2011; Warren, Pollock, White, Day, et al., 2011; Warren, White, Day, & Pollock, 2011) provide exemplars of the application of the framework. We also applied the framework in evaluations of the N.Z. National Shared Care Planning pilot for long-term condition management (National Institute for Health Innovation, 2013; Warren, Gu, & Humphrey, 2012; Warren, Humphrey, & Gu, 2011) and the Canterbury electronic Shared Care Record View project (Gu, Humphrey, Warren, & Wilson, 2014).

6.4.2 Implications

The key contribution of the evaluation against the benefits framework should be to indicate whether the innovation is one that should be adopted broadly. To warrant recommendation for emulation the innovation should be free of "red flags" — this includes being free of evidence of net harm to patients, and having no major negative impact on the health workforce. Beyond this, the innovation must show a clear case for some benefit that is sufficiently compelling to warrant the cost and disruption of adopting the innovation.

Health workforce is a particular challenge for many healthcare systems; certainly it is for New Zealand, where we face a "complex demand-supply-affordability mismatch" (Gorman, 2010). As such, benefits that tie directly back to effective use of health workforce will be particularly compelling. If an innovation allows more to be done (at the same quality) with the same number of health-

care workers, or allows doing better with the same number of healthcare workers, then it is compelling. An innovation that empowers and satisfies healthcare workers may also be compelling due to its ability to retain those workers. And an innovation that lets workers "practice at the top of their licence" (Wagner, 2011) will get the most out of our limited health workforce and engender their satisfaction while doing so. In some cases this may involve changing care delivery patterns in accordance with evidence-based medicine such that use of particular services or procedures is reduced (e.g., shifting service from hospital-based specialist care to the community). Such changes should be detectable from the transactional records of health information systems.

6.5 Summary

A pragmatic evaluation framework has been recommended for projects involving innovative use of health IT. The framework recommends using both qualitative and quantitative data. It emphasizes capture of a broad range of stakeholder perspectives and multidimensional evaluation on criteria related to process and culture, as well as outcome and IT system integrity. It also recommends underpinning quantitative analysis with the transactional data from the health IT systems themselves.

The recommended approach is iterative and Action Research (AR) oriented. Evaluation should be integral to implementation. It should begin, if possible, before the new technology is introduced into the health workflow and be planned for along with the planning of the implementation itself. Evaluation findings should be used to help refine the implementation and to evoke further user feedback. Dissemination of the findings is integral and should reach all stakeholders considering uptake of similar technology.

References

Baum, F., MacDougall, C., & Smith, D. (2006). Participatory action research. *Journal of Epidemiology & Community Health, 60*(10), 854–857. doi: 10.1136/jech.2004.028662

Cooper, A., Reinmann, R. M., & Cronin, D. (2007). *About face 3: the essentials of interaction design* (3rd ed.). New York: John Wiley & Sons.

Day, K., Gu, Y., Warren, J., White, S., & Pollock, M. (2011). *National eReferral evaluation: Findings for Northland District Health Board*. Wellington: Ministry of Health.

DeLone, W. H., & McLean, E. R. (2003). The DeLone and McLean model of information system success: A ten-year update. *Journal of Management Information Systems, 19*(4), 9–30.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago: Aldine Publishing Company.

Gorman, D. (2010). The future disposition of the New Zealand medical workforce. *New Zealand Medical Journal, 123*(1315), 6–8.

Greenhalgh, T., & Russell, J. (2010). Why do evaluations of eHealth programs fail? An alternative set of guiding principles. *Public Library of Science Medicine, 7*(11), e1000360. doi: 10.1371/journal.pmed.1000360

Gu, Y., Day, K., Humphrey, G., Warren, J., & Pollock, M. (2012). *National eReferral evaluation: Findings for Waikato District Health Board*. Wellington: Ministry of Health. Retrieved from http://api.ning.com/files/FrmWaEpjYR4bwaOAyZTeTNjQ65183Caca4ZDkjOg*IDLoICJEVdtCzbAoK3XRUY27MYXWeF7536QOsGHiMfkcJG1cRQu*StU/Waikato.eReferral.evaluation.pdf

Gu, Y., Humphrey, G., Warren, J., & Wilson, M. (2014, December). *Facilitating information access across healthcare settings — A case study of the eShared Care Record View project in Canterbury, New Zealand.* Paper presented at the 35th International Conference on Information Systems, Auckland, New Zealand.

Gu, Y., Warren, J., & Day, K. (2011). Unleashing the power of human genetic variation knowledge: New Zealand stakeholder perspectives. *Genetics in Medicine, 13*(1), 26–38. doi: 10.1097/GIM.0b013e3181f9648a

IT Health Board. (2010). *National health IT plan: Enabling an integrated healthcare model.* Wellington: National Health IT Board.

Klein, H., & Myers, M. (1999). A set of principles for conducting and evaluating interpretive field studies in information systems. *Management Information Systems Quarterly, 23*(1), 67–93.

Lau, F., Hagens, S., & Muttitt, S. (2007). A proposed benefits evaluation framework for health information systems in Canada. *Healthcare Quarterly, 10*(1), 112–118.

Lau, F., Kuziemsky, C., Price, M., & Gardner, J. (2010). A review on systematic reviews of health information system studies. *Journal of the American Medical Informatics Association, 17*(6), 637–645. doi: 10.1136/jamia.2010.004838

Macaulay, A. C., Commanda, L. E., Freeman, W. L., Gibson, N., McCabe, M. L., Robbins, C. M., & Twohig, P. L. (1999). Participatory research maximises community and lay involvement. (Report for the North American Primary Care Research Group). *British Medical Journal, 319*(7212), 774–778.

McNiff, J., & Whitehead, J. (2002). *Action research: principles and practice* (2nd ed.). London: Routledge.

National Institute for Health Innovation. (2013, May). Shared Care Planning evaluation Phase 2: Final report. Retrieved from http://www.sharedcareplan.co.nz/Portals/0/documents/News-and-Publications/Shared%20Care%20Planning%20Evaluation%20Report%20-%20Final2x.pdf

Neuman, W. L. (2003). *Social research methods: Qualitative and quantitative approach* (5th ed.). Boston: Pearson Education.

Stringer, E. T. (1999). *Action research* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Wagner, E. H. (2011). *Better care for chronically ill people* (presentation to the Australasian Long-Term Conditions Conference, April 7, 2011). Retrieved from http://www.healthnavigator.org.nz/conference/presentations

Warren, J., Gu, Y., Day, K., Pollock, M., & White, S. (2012). Approach to health innovation projects: Learnings from eReferrals. *Health Care and Informatics Review Online, 16*(2), 17–23.

Warren, J., Gu, Y., & Humphrey, G. (2012). *Usage analysis of a shared care planning system.* Paper presented at the The AMIA 2012 Annual Symposium, Chicago.

Warren, J., Humphrey, G., & Gu, Y. (2011). *National Shared Care Planning Programme (NSCPP) evaluation: Findings for Phase 0 & Phase 1*. Wellington: Ministry of Health.Retrieved from http://www.sharedcareplan.co.nz/Portals/0/documents/News-and-Publications/F.%20NSCPP_evaluation_report%2020111207%202.pdf

Warren, J., Pollock, M., White, S., & Day, K. (2011). *Health IT evaluation framework*. Wellington: Ministry of Health.

Warren, J., Pollock, M., White, S., Day, K., & Gu, Y. (2011). *A framework for health IT evaluation.* Paper presented at the Health Informatics New Zealand 10th Annual Conference and Exhibition, Auckland.

Warren, J., White, S., Day, K., & Pollock, M. (2011). *National eReferral evaluation: Findings for Hutt Valley District Health Board*. Wellington: Ministry of Health.

Westbrook, J. I., Braithwaite, J., Georgiou, A., Ampt, A., Creswick, N., Coiera, E., & Iedema, R. (2007). Multimethod evaluation of information and communication technologies in health in the context of wicked problems and sociotechnical theory. *Journal of the American Medical Informatics Association, 14*(6), 746–755. doi: 10.1197/jamia.M2462.

Chapter 7

Holistic eHealth Value Framework

Francis Lau, Morgan Price

7.1 Introduction

Canadian jurisdictions have been investing in health information technology (HIT) as one strategy to address healthcare sustainability. Investments have included the migration to electronic patient records, and the automation of service delivery to improve the efficiency, access and quality of care provided. In this context, eHealth emerged over 10 years ago as a shared priority for the federal, provincial and territorial jurisdictions in their health care renewal effort. To date, the federal government has invested over $2 billion in Canada Health Infoway (Infoway) through incremental and targeted funding. Provinces and territories have also invested in the cost sharing of eHealth projects. Progress has been made towards achieving the eHealth vision. Examples are: (a) the adoption of pan-Canadian approaches among provinces and territories in the planning and development of common EHR architectures and standards; (b) the creation of jurisdictional registries, such as patient and provider registries; and (c) the creation of jurisdictional repositories of patient data, such as imaging, lab and drug information systems.

Yet there is conflicting evidence on eHealth benefit. Some reports suggest strong benefit while others showed few to no benefits in spite of the eHealth investments made. For example, in their 2009-2010 performance audit reports, the Auditor General of Canada and six provincial auditors' offices raised questions on whether there was sufficient "value for money" from the EHR investments (e.g., Office of the Auditor General of Canada [OAG], 2010). In light of the investments made, an effort is needed to make sense of the evidence on eHealth benefits. To do so, we created a high-level conceptual eHealth Value Framework as an organizing scheme to examine the current evidence on

Canadian eHealth value, and the underlying reasons for the conflicting evidence so that future eHealth investment and work is better informed.

This chapter describes a proposed holistic eHealth Value Framework to make sense of the value of eHealth systems in the Canadian setting. The chapter contains an overview of this framework, its use in a Canadian literature review on eHealth value, and implications on policy and practice.

7.2 A Sense-making Scheme for eHealth Value

The proposed holistic eHealth Value Framework is described in this section in terms of its conceptual foundations and the respective framework dimensions.

7.2.1 Conceptual Foundations

The eHealth Value Framework incorporates several foundational frameworks and models from the literature. The underpinnings of this framework are the following: the Infoway Benefits Evaluation (BE) Framework (Lau, Hagens, & Muttitt, 2007); the Clinical Adoption Framework (Lau, Price, & Keshavjee, 2011); the Clinical Adoption and Maturity Model (eHealth Observatory, 2013); Canada's Health Informatics Association [COACH] Canadian EMR Adoption and Maturity Model (COACH, 2013); the HIMSS EMR Adoption Model (HIMSS Analytics, 2014); Meaningful Use Criteria (Blumenthal & Tavenner, 2010); and the Information Systems Business Value Model (Schryen, 2013). By combining features of these models, this framework provides a comprehensive view of eHealth, incorporating, for example, the EHR and its value.

7.2.2 Value Framework Dimensions

The *eHealth Value Framework for Clinical Adoption and Meaningful Use* (hereafter referred to as the *eHealth Value Framework*) describes how the value of eHealth components, such as an EHR, is influenced by the dynamic interactions of a complex set of contextual factors at the micro, meso, and macro adoption levels. The outcomes of these interactions are complex. The realized benefits (i.e., the value of an EHR) depend on the type of investment made, the system being adopted, the contextual factors involved, the way these factors interact with each other, and the time for the system to reach a balanced state. Depending on the adjustments made to the system and the adoption factors along the way, the behaviour of this system and its value may change over time.

Specifically, there are four interrelated dimensions that can be used to explain the benefits of EHRs. They are: *Investment*, *Adoption*, *Value*, and *Time*. Each is made up of a set of contextual factors that interact dynamically over time to produce specific EHR impacts and benefits (see Figure 7.1). These dimensions are described next.

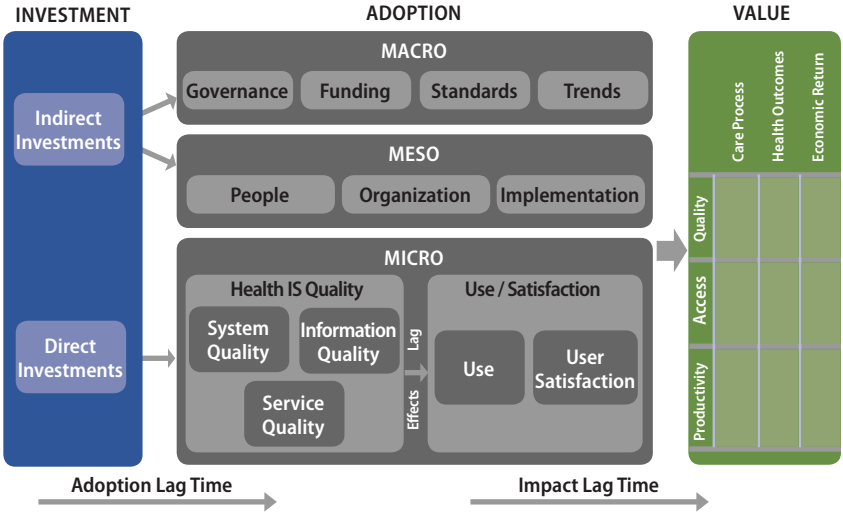


Figure 7.1. A proposed holistic eHealth value framework for clinical adoption and meaningful use.

7.2.3 Investment

Investments can be made directly towards achieving EHR adoption or indirectly to influence larger contextual factors that impact adoption.

7.2.4 Adoption

Adoption can be considered at a micro level, consistent with the Infoway BE Framework. It also has contextual factors at the meso and macro levels, ranging from people and organizational structures to larger standards, funding structures, and pieces of legislation.

- Micro – The quality of the system and its use can influence the intended benefits. The technology, information, and support services provided can influence how the system performs. This can impact the actual or intended use of the system and user satisfaction. If a system does not support certain functionality (e.g., system quality), or is not used appropriately or as intended, value is not likely to be seen.

- Meso – People, organization, and implementation processes can influence the intended benefits of the system. People refer to those individuals/groups that are the intended users, their personal characteristics and expectations, and their roles and responsibilities. Organizations have strategies, cultures, structures, processes, and

info/infrastructures. Implementation covers the system's life cycle stages, its deployment planning/execution process, and the system's fit for purpose.

- Macro – Governance, funding, standards and trends can influence the benefits. Governance refers to legislation, policies and accountability. Funding includes remunerations, incentives and added values for the system. Standards include HIT, performance, and practice standards. Trends cover the general public, political and economic investment climates toward EHR systems.

7.2.5 Value

Value of EHR is defined as the intended benefits from the clinical adoption and meaningful use of the EHR system. Value can be in the form of improved care quality, better access, and increased productivity affecting care processes, health outcomes, and economic return. It can be measured by different methods and at various times in relation to adoption.

7.2.6 Lag Time

There is an acknowledged lag time to implement and realize benefits from EHR adoption. Lag effects occur as EHR systems become incorporated into practice, where adoption factors at the micro, meso and macro levels can all impact lag time until benefits from the adoption are evident.

7.3 Framework Use and Implications

This section describes the use of the eHealth Value Framework to make sense of eHealth benefit with respect to a literature review undertaken in 2014 on a set of Canadian eHealth evaluation studies published between 2009 and 2013. Three Canadian literature sources were included: 12 Infoway co-funded benefits evaluation studies; 25 primary studies in peer-reviewed journals; and one federal government auditor's report. The systems evaluated were EHRs, drug information systems (DIS), lab information systems, diagnostic imaging (DI/PACS), ePrescribing, computerized provider order entries (CPOEs), provincial drug viewers, and physician office EMRs (Lau, Price, & Bassi, 2014).

7.3.1 Use

The eHealth Value Framework was applied to organize the review findings; eHealth benefit was examined through the value dimensions of care process, health outcomes, and economic return. Factors that influence adoption were examined at the micro, meso and macro level of the adoption dimension. Of the 38 Canadian studies reviewed, 21 had reported benefit findings, 29 had reported adoption factors, and 21 had evaluated and reported on the adoption factors. Of the 21 studies on benefit, there was a combination of positive, mixed,

neutral and negative benefits reported (see Figure 7.2). Overall, there appears to be a small but growing body of evidence on the adoption, impact and value of eHealth systems in Canada. These benefits are summarized below according to the value dimension of the framework.

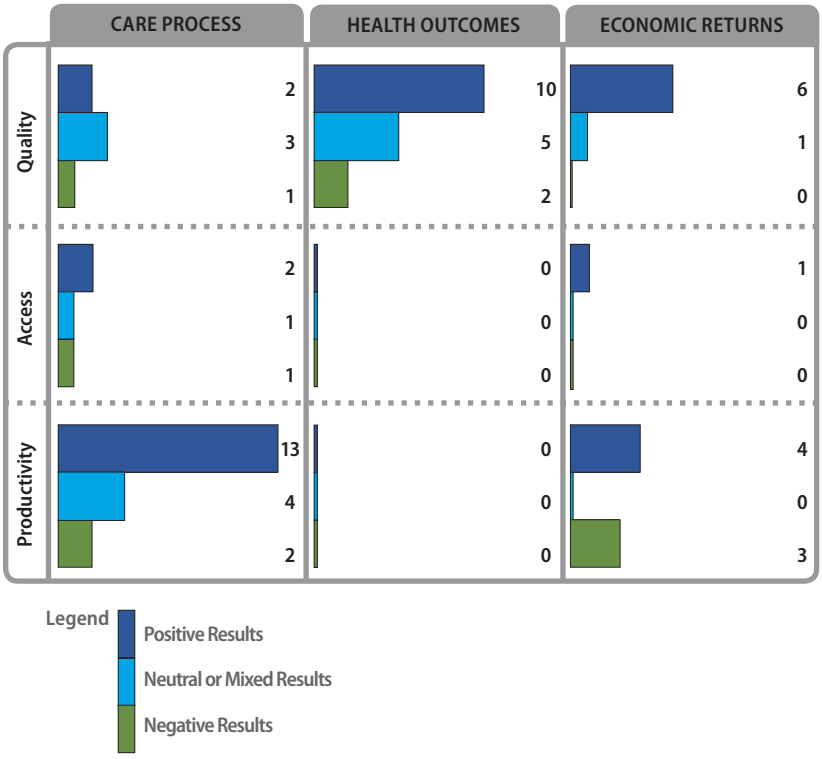


Figure 7.2. Summary of eHealth value findings from Canadian studies.

1 Care Process – Most of the studies reported benefits in care process (actual or perceived improvements). These care processes involved activities that could improve patient safety (Tamblyn et al., 2010; Geffen, 2013), guideline compliance (Holbrook et al., 2009; PricewaterhouseCoopers [PWC], 2013; Geffen, 2013), patient/provider access to services (Geffen, 2013; Prairie Research Associates [PRA], 2012), patient-provider interaction (Holbrook et al., 2009; Centre for Research in Healthcare Engineering [CRHE], 2011), productivity/efficiency (Prince Edward Island [P.E.I.] Department of Health and Wellness, 2010; Paré et al., 2013; CRHE, 2011; Lapointe et al., 2012; Syed et al., 2013), and care coordination (Paré et al., 2013; PWC, 2013; Lau, Partridge, Randhawa,

& Bowen, 2013). There were also some negative impacts which included poor EMR data quality that affected drug-allergy detection (Lau et al., 2013), perceived inability of the EMR to facilitate decision support (Paré et al., 2013), increased pharmacist callback in ePrescribing (Dainty, Adhikari, Kiss, Quan, & Zwarenstein, 2011), and reduced ability of a DIS to coordinate care and share information (P.E.I. Department of Health and Wellness, 2010).

2 Health Outcomes – The overall evidence on health outcome benefits is smaller and is more mixed. Two controlled DIS studies reported improved patient safety with reduced inappropriate medications (Dormuth, Miller, Huang, Mamdani, & Juurlink, 2012) and errors (Fernandes et al., 2011), while a third study reported low accuracy of selected medications in a provincial medication dispensing repository (Price, Bowen, Lau, Kitson, & Bardal, 2012). On the other hand, two descriptive studies reported user expectations of improved compliance and reduced adverse events with full DIS adoption and use. For EMR, Holbrook et al. (2009) reported improved A1c and blood pressure control levels, while Paré et al. (2013), PWC (2013) and Physician Information Technology Office [PITO] (2013) all reported expectations of improved safety from the EMR. At the same time, PITO (2013) reported that less than 25% of physicians believed EMR could enhance patient-physician relationships and Paré et al. (2013) reported few physicians believed EMR could improve screening. For ePrescribing and CPOE there were no improved outcomes in patient safety reported (Tamblyn et al., 2010; Dainty et al., 2011; Lee et al., 2010).

3 Economic Return – The overall evidence on economic return is also mixed. For EMR, O'Reilly, Holbrook, Blackhouse, Troyan, and Goeree, (2012) reported a positive return on diabetes care from Holbrook et al.'s original 2009 RCT study that showed an improved health outcome of 0.0117 quality-adjusted life years with an incremental cost-effectiveness ratio of $160,845 per quality-adjusted life year. PRA (2012) reported mixed returns where the screening of breast and colorectal cancers was cost-effective but not in cervical cancer. In Paré et al.'s (2013) survey less than 25% of Quebec physicians reported direct linkage between the EMR and financial health of their clinics. The PITO (2013) survey also reported that less than 25% of British Columbia physicians believed EMR could reduce overall office expenses. The PWC study (2013) estimated the combined economic return from productivity and care quality improvements to be $300 million per year with full EMR adoption and use. For DI/PACS, MacDonald and Neville (2010) reported a

negative return of the P.E.I. PACS system from their cost-benefit analysis with an increased cost per exam, which was estimated to take six years to amortize with the higher cost. On the other hand, Geffen (2013) estimated a positive return of $89.8 million per year in DI/PACS based on its full adoption and optimal use in B.C. For DIS, both the Deloitte (2010) and Geffen (2013) studies estimated positive returns in excess of $435 million and $200 million per year nationally and in B.C., respectively. Their predictions are based on full adoption and use of the systems.

7.3.2 Clinical Adoption of eHealth Systems

To better understand why the value of eHealth is not consistently being realized, it is prudent to consider the contextual factors surrounding adoption that influence these findings. Put differently, the value derived from eHealth is dependent on these contextual factors, which affect the extent of system adoption that takes place in an organization. Not all studies addressed issues of adoption to explain their findings; 29 of the Canadian studies did report contextual factors for adoption. The identified factors were mapped to the adoption dimension of the eHealth Value Framework, with specific examples in each category. They are explained below and summarized in Table 7.1.

1 Micro level – The design of the system in terms of its functionality, usability and technical performance had a major influence on how it was perceived and used, which in turn influenced the actual benefits. For instance, the P.E.I. DIS (P.E.I. Department of Health and Wellness, 2010) users had mixed perceptions on the system's ease of use, functionality, speed, downtime and security that influenced their use and satisfaction. The quality of the clinical data in terms of accuracy, completeness and relevance influenced its clinical utility. The actual system use and its ability to assist in decision-making, data exchange and secondary analysis also influenced its perceived usefulness. For instance, seven of the EMR studies involved the development and validation of algorithms to identify patients with specific conditions (Tu et al., 2010a; Tu et al., 2010b; Tu et al., 2011; Harris et al., 2010; Poissant, Taylor, Huang, & Tamblyn, 2010; Roshanov, Gerstein, Hunt, Sebaldt, & Haynes, 2013), generate quality indicators (Burge, Lawson, Van Aarsen, & Putnam, 2013), and conduct secondary analyses (Tolar & Balka, 2011). The type and extent of user training and support also influenced adoption. Shachak, Montgomery, Tu, Jadad, and Lemieux-Charles (2013) identified different types of end user support sources, knowledge and activities needed to improve EMR use over time.

2 Meso level – For people, the level of user competence, experience and motivation, the capability of the support staff, and the availability of mentors all influenced adoption. For instance, Lapointe et al. (2012) found providers had varying abilities in performing EMR queries to engage in reflective practice on their patient populations. The end user support scheme identified by Shachak et al. (2012) directly influenced the confidence and capabilities of the users and support staff. Even after implementation, time was still needed for staff to learn the system, as was reported by Terry, Brown, Denomme, Thind, and Stewart (2012) with respect to users of EMRs that had been implemented for two years. For organizations, having management commitment and support, realistic workload, expectations and budgets, and an interoperable infrastructure influenced adoption. These factors were reported by McGinn et al. (2012) in their Dephi study on successful implementation strategies with representative EHR user groups. For implementation, the ability to manage the project timeline, resources and activities, and to engage providers all had major influences on successful adoption. An example was the health information exchange (HIE) study reported by Sicotte and Paré (2010), where the implementation efforts had major influences on the success or failure of two HIE systems. The Auditor General's report (OAG, 2010) raised concerns with EHR implementation initiatives in terms of insufficient planning, governance, monitoring and public reporting that led to unclear value for money.

3 Macro level – One study addressed the standards, funding, and policy aspects of the Canadian eHealth plan to adopt an interoperable EHR (Rozenblum et al., 2011). Rozenblum and colleagues acknowledged Canada's national eHealth standards, EHR funding, registries and DI/PACS as tangible achievements over the past 10 years. Yet these authors felt the Canadian plan fell short of having a coordinated eHealth policy, active clinician engagement, a focus on regional interoperability, a flexible EHR blueprint, and a business case to justify the value of an EHR. As recommendations, their study called for an eHealth policy that is tightly aligned with major health reform efforts, a bottom-up approach by placing clinical needs first with active clinician and patient engagements, coordinated investments in EMRs to fill the missing gap, and financial incentives on health outcomes that can be realized with EHRs. Similarly, McGinn et al. (2012) and PITO (2013) suggested physician reimbursement and incentives as ways to encourage EMR adoption. Burge et al. (2013), Holbrook et al. (2009) and Eguale, Winslade, Hanley, Buckeridge, and Tamblyn (2010) all emphasized

the need for data standards to improve interoperability. Note that Infoway received additional funding in 2010 to expand their scope to include support for physician EMRs, which include clinician engagement through such efforts as the Clinician Peer Support Network (Infoway, 2013).

Table 7.1

Summary of Adoption Factors that Influence eHealth Values from Canadian Studies

	Adoption/Impact Factors	Canadian Studies
Micro Level		
System	Functionality/features; System design; Usability; Technical issues; Privacy and security concerns.	Dainty et al. (2011); Eguale et al. (2010); Poissant et al. (2010); Price et al. (2012); PRA (2012); Paré et al. (2013); Lapointe et al. (2012); Deloitte (2010); McGinn et al. (2012); Paterson et al. (2010); G. Braha & Associates (2012); Holbrook et al. (2009)
Information	Database completeness; Structured data; Data quality; Volume of data; Enhanced information; Information capture.	Eguale et al. (2010); Tu et al. (2011); Lau et al. (2013); Deloitte (2010); MacDonald and Neville (2010); Fernandes et al. (2011); Burge et al. (2013); Tu et al. (2010a; 2010b)
Service	Resources and support; Training; Learning curve; Support personnel; Communication/information to end- users; Infrastructure support; Learning space.	McGinn et al. (2012); Mensink and Paterson (2010); P.E.I. (2010); MacDonald and Neville (2010); PITO (2013); Lapointe et al. (2012); Lau et al. (2013); Paré et al. (2013); Deloitte (2010); Shachak et al. (2012)
Use	Use/variability in use; Perceived usefulness.	Paterson et al. (2010); Terry et al. (2012); Tolar and Balka (2011); McGinn et al. (2012)
Satisfaction	Familiarity/confidence with system; Interaction with computer; Learning to use system; Perceived ease of use.	McGinn et al. (2012); Terry et al. (2012); Eguale et al. (2010); Paré et al. (2013)
Meso Level		
People	Client/user population; Individual user behaviours; Champions/super users; Confidence with computers; User expectations; Roles and responsibilities; Meaningful engagement of clinicians.	Shachak et al. (2012); MacDonald and Neville (2010); Deloitte (2010); Dainty et al. (2011); G Braha and Associates (2012); Terry et al. (2012); Mensink and Paterson (2010); Lau et al. (2013); Rozenblum et al. (2011)
Organization	Business requirements/planning; Implementation strategy/change management; Vision/long term planning; Participation of end-users in implementation strategy; Organizational readiness; Defined value; Commitment; Individual workplace type; Management; Communication; Leadership; Workplace size; Physician salary; Scanning in documents; Internet connectivity; Current/prior technology in use; National infrastructure; Entry of information; Interoperability/connectivity; Cost issues/benefit.	Lau et al. (2013); PEI (2010); Paterson et al. (2010); Sicotte and Paré (2010); G. Braha and Associates (2012); Mensink and Paterson (2010); PITO (2013); Deloitte (2010); McGinn et al. (2012); PRA (2012); Dainty et al. (2011); MacDonald and Neville (2010); Shachak et al. (2012); Lapointe et al. (2012); Terry et al. (2012); Holbrook et al. (2009); Rozenblum et al. (2011); Paré et al. (2013); CRHE (2011); O'Reilly et al. (2012); PITO (2013); OAG (2010)

Meso Level		
Implementation	Uptake; Loss of productivity; Choice of system; Response to risks; Implementation team; Experience in IT project management; Implementation effort; Complexity; Confidence in system developer or vendor/communication with vendor; Workload/workflow; User vs. vendor needs; Change in tasks.	P.E.I. (2010); G. Braha and Associates (2012); Deloitte (2010); McGinn et al. (2012); Sicotte and Paré (2010); Lau et al. (2013); CRHE (2011); Paré et al. (2013); PITO (2013); Paterson et al. (2010); Shachak et al. (2012); MacDonald and Neville (2010); PRA (2012); OAG (2010)
Macro Level		
Standards	Standards for interoperability; Standards for data structure and extraction; Security standards; Standardized coding; Quality standards; Clinical best practices; Standardization of data entry.	Deloitte (2010); Rozenblum et al. (2011); Burge et al. (2013); Holbrook et al. (2009); P.E.I. (2010); Tu et al. (2010b); Eguale et al. (2010)
Funding/ Incentives	Physician reimbursement; Motivation; Secondary uses; Education; Gated funding; Peer competition; Financial incentives.	McGinn et al. (2012); PITO (2013); Paterson et al. (2010); PRA (2012); Deloitte (2010); Lapointe et al. (2012); Rozenblum et al. (2011); PRA (2012)
Legislation/ Policy/ Governance	Legislation to guide use (e.g., electronic signatures); National policy for effective adoption strategies; Alignment with healthcare transformation agenda; System certification; Flexible blueprint adaptive to feedback from implementation; Framework for collaboration across jurisdictions; Coordinated national leadership and investment; Accountability through public reporting.	Dainty et al. (2011); Deloitte (2010); Rozenblum et al. (2011); Paré et al. (2013); OAG (2010)
Time lags		
Adoption	Lack of time; Time to integrate system into daily practice; Project time.	McGinn et al. (2012); Tolar and Balka (2011); Sicotte and Paré (2010); PITO (2013); Deloitte (2010)
Impact	Short follow-up time; Early stage of implementation; Lag time for benefits.	Holbrook et al. (2009); Lau et al. (2013); P.E.I. (2010); Paré et al. (2013); Deloitte (2010)

7.3.3 Evaluations of Clinical Adoption Factors

In addition to mentioning the contextual factors that may have been facilitators or barriers to achieving value in care, process and return, 21 studies actually evaluated some influencing adoption factors themselves. For example, in the three papers by Tu and colleagues (2010a, 2010b, 2011), the primary focus was on the content of the EMRs and its ability to help identify patient populations. While this may not be an example of a health value outcome, it is an example of an information content measure that contributes to care provision. The rationale is that the value of the EMR is dependent on the quality of the data. If data quality is lacking, then value at the health outcome level will be impacted. Therefore, looking at individual factors from an evaluation perspective may also help to make sense of the evidence. Simply having the factor present — for example, training for end users — does not ensure successful outcomes. The findings for factors examined in the Canadian studies are summarized in Figure 7.3.

Figure 7.3. Summary of adoption factors assessed in micro, meso, and macro categories. There is a considerable focus on micro factors and it was challenging to find assessment of macro level factors.

7.4 Implications

The current evidence on Canadian eHealth benefits is confusing and difficult to interpret even for the experienced eHealth researcher and practitioner. There are four types of issues that should be considered when navigating the eHealth benefits landscape. These are the definition of eHealth, one's views or perception of the eHealth system, the methods used to study benefits, and system adoption

that can influence eHealth benefits. These issues and their implications for healthcare organizations are discussed below.

- *Definitions* – The field of eHealth is replete with jargon, acronyms and conflicting descriptions. For instance, eHealth refers to the application of health information and communication technology or ICT in health. It is a term often seen in the Canadian and European literature. On the other hand, health information technology or HIT describes the use of ICT in health especially in the United States. The terms EHR and EMR can have different meanings depending on the countries in which they are used. In the U.S., EHR and EMR are used interchangeably to mean electronic records that store patient data in healthcare organizations. However, in Canada EMR refers specifically to electronic patient records in a physician's office. The term EHR can also be ambiguous. According to the Institute of Medicine, an EHR has four core functions of health information and data, order entry (i.e., CPOE), results management, and decision support (Blumenthal et al., 2006). Sometimes it may also include patient support, electronic communication and reporting, and population health management. Even CPOE can be ambiguous as it may or may not include decision support functions. The challenge with eHealth definitions, then, is that there are often implicit, multiple and conflicting meanings. The Canadian eHealth literature is no exception. Thus, when reviewing the Canadian evidence on eHealth benefits one needs to understand what system and/or function is involved, how it is defined and where it is used.

- *Views or perceptions* – The type of eHealth system and function being evaluated, the care setting involved, and the focus of the evaluation are important considerations that influence how the system is viewed or perceived by different stakeholders as to its intentions, roles and values. Most evaluation studies would identify the eHealth system and/or function being investigated, such as an EHR with CDS and/or CPOE. The care setting can influence how a system is adopted since it embodies the type of care and organizational practices being provided. The focus is the clinical area being evaluated and the benefit expected, such as medication management with CPOE to reduce errors. The challenge with eHealth views as articulated in these studies, then, is that the descriptions of the system, setting and focus are often incomplete in the evaluation write-up, which makes it difficult to determine the relevance of the findings to the local setting. For example, in studies of CPOE with alerts, it is often unclear how they are generated and to whom, and whether a response is required. For a setting such

as a primary care clinic it is often unclear whether the actual site is a hospital outpatient department or a stand-alone community-based practice. For focus, some studies include such a multitude of benefit measures that it can be difficult to decide if the system has led to overall benefit. The Canadian eHealth studies face the same challenge of having to tease out such detail to determine the relevance and applicability of the findings.

- *Methods of study* – There is a plethora of scientific, psychosocial and business methods used to evaluate eHealth benefits. At one end of the spectrum are such experimental methods as the randomized control trial (RCT) used to compare two or more groups for notable changes from the implementation of an eHealth system as the intervention. At the other end is the descriptive method used to explore and understand the interactions between an eHealth system and its users. The choice of benefit measures selected, the type of data collected and the analytical method used can all affect the study results. In contrast to controlled studies that strive for statistical and clinical significance in the outcome measures, descriptive studies offer explanations of the observed changes as they unfold. There are also economic evaluation methods that examine the relationships between the costs and return of an investment, and simulation methods that model changes based on a set of input parameters and analytical algorithms. The challenge, then, is that one needs to know the principles and rigour of different methods in order to plan, execute, and appraise eHealth benefits evaluation studies. The Canadian eHealth evidence identified in this chapter has been derived from different approaches such as RCTs, descriptive studies and simulation methods. The quality of these studies varies depending on the rigour of the design/method used. The different outcome measures used has made it difficult to aggregate the findings. Finally, timing of studies in relation to adoption and use will influence benefits, which may or may not be seen.

- *System adoption* – There are mixed and even conflicting results from evaluation studies on eHealth benefits. To understand these differences one has to appreciate the context surrounding the implementation, use and impacts of eHealth systems in organizations. The success of an eHealth system in producing the expected benefits is dependent on many contextual factors. Examples are the usability of the system involved, prior experience of its users, the training and support available, the organizational culture and commitment toward eHealth and the system, how well the implemen-

tation process is managed, the funding and incentives in place and the overall expectations. The contextual factors are described in detail under the investment, micro, meso, macro and value dimensions of the proposed eHealth Value Framework presented in this chapter. These contextual factors apply equally well to the Canadian eHealth systems being evaluated. The challenge, then, is whether the level of detail provided in the evaluation write-up is sufficient, and whether it can explain why the system had worked or not, and if not, what could be done to achieve the benefits.

7.5 Summary

This chapter introduced the holistic eHealth Value Framework to make sense of eHealth value in the Canadian setting. This framework is made up of four dimensions of investment, adoption, value and time lag. It was applied in a review of Canadian literature on eHealth evaluation studies to examine eHealth value within the Canadian context. The framework helped to make sense of the conflicting evidence found in the literature on eHealth benefits in Canada.

References

Blumenthal, D., DesRoches, C., Donelan, K., Ferris, T., Jha, A., Kaushal, R., … Shield, A. (2006). *Health information technology in the United States: The information base for progress.* Princeton, NJ: Robert Wood Johnson Foundation.

Blumenthal, D., & Tavenner, M. (2010). The "meaningful use" regulation for electronic health records. *The New England Journal of Medicine*, *363*(6), 501–504.

Burge, F., Lawson, B., Van Aarsen, K., & Putnam, W. (2013). Assessing the feasibility of extracting clinical information to create quality indicators from primary healthcare practice EMRs. *Healthcare Quarterly*, *16*(3), 34–41.

Canada's Health Informatics Association [COACH]. (2013). *Canadian EMR adoption and maturity model*. Toronto: Author. Retrieved from http://www.coachorg.com/en/resourcecentre/Green_White-Papers.asp

Centre for Research in Healthcare Engineering [CRHE]. (2011). *EMR integrated labs workflow evaluation.* Toronto: University of Toronto. Retrieved from https://www.infoway-inforoute.ca/index.php/resources/reports

Dainty, K. N., Adhikari, N. K. J., Kiss, A., Quan, S., & Zwarenstein, M. (2011). Electronic prescribing in an ambulatory care setting: a cluster randomized trial. *Journal of Evaluation in Clinical Practice*, *18*(4), 761–767.

Deloitte. (2010). *National impacts of generation 2 drug information systems technical report*. Toronto: Canada Health Infoway. Retrieved from https://www.infoway-inforoute.ca/index.php/resources/reports

Dormuth, C., Miller, T. A., Huang, A., Mamdani, M. M., & Juurlink, D. N. (2012). Effect of a centralized prescription network on inappropriate prescriptions for opioid analgesics and benzodiazepines. *Canadian Medical Informatics Journal*, *184*(16), E852–E856.

Eguale, T., Winslade, N., Hanley, J. A., Buckeridge, D. L., & Tamblyn, R. (2010). Enhancing pharmacosurveillance with systematic collection of treatment indication in electronic prescribing: a validation study in Canada. *Drug Safety*, *33*(7), 559–567.

eHealth Observatory. (2013). *Clinical adoption maturity model*. Victoria, BC: University of Victoria. Retrieved from http://ehealth.uvic.ca/methodology/models/CMM.php

Fernandes, O. A., Etchells, E. E., Lee, A. W., Siu, V., Bell, C., Wong, G., … Department of Medicine, McMaster University. (2011). What is the impact of a centralized provincial drug profile viewer on the quality and efficiency of patient admission medication reconciliation? A randomized controlled trial. *Canadian Journal of Hospital Pharmacy, 64*(1), 85.

G. Braha & Associates Ltd. (2012). *eChart Manitoba release 1: a benefits evaluation*. Winnipeg, MB: Author. Retrieved from https://www.infoway-inforoute.ca/index.php/resources/reports

Geffen, M. (2013). *British Columbia eHealth benefits estimates*. Stamford, CT: Gartner Consulting. Retrieved from https://www.infoway-inforoute.ca/index.php/resources/reports

Harris, S. B., Glazier, R. H., Tompkins, J. W., Wilton, A. S., Chevendra, V., Stewart, M. A., & Thind, A. (2010). Investigating concordance in diabetes diagnosis between primary care charts (electronic medical records) and health administrative data: a retrospective cohort study. *BMC Health Services Research*, *10*, 347.

HIMSS Analytics. (2014). *EMR adoption model*. Chicago: Author. Retrieved from http://www.himssanalytics.org/home/index.aspx

Holbrook, A., Thabane, L., Keshavjee, K., Dolovich, L., Bernstein, B., Chan, D., … Gerstein, H. (2009). Individualized electronic decision support and reminders to improve diabetes care in the community: COMPETE II randomized trial. *Canadian Medical Association Journal*, *181*(1/2), 37–44.

Infoway. (2013). *Jurisdictional peer-to-peer networks*. Toronto: Author. Retrieved from https://www.infoway-inforoute.ca/en/communities/pan-canadian-clinician-peer-network/jurisdictional-peer-to-peer-networks

Lapointe, L., Hughes, J., Simkus, R., Lortie, M., Sanche, S., & Law, S. (2012). *The population health management challenge final report.* Montreal: St. Mary's Research Centre. Retrieved from https://www.infoway-inforoute.ca/index.php/resources/reports

Lau, F., Hagens, S., & Muttitt, S. (2007). A proposed benefits evaluation framework for health information systems in Canada. *Healthcare Quarterly, 10*(1), 112–118.

Lau, F., Partridge, C., Randhawa, G., & Bowen, M. (2013). Applying the clinical adoption framework to evaluate the impact of an ambulatory electronic medical record. *Studies in Health Technology & Informatics*, *183*, 15–20.

Lau, F., Price, M., & Bassi, J. (2014). *Making sense of eHealth benefits and their policy implications in Canada*. Unpublished discussion paper.

Lau, F., Price, M., & Keshavjee, K. (2011). From benefits evaluation to clinical adoption: Making sense of health information system success. *Healthcare Quarterly, 14*(1), 39–45.

Lee, J. Y., Leblanc, K., Fernandes, O. A., Huh, J., Wong, G. G., Hamandi, B., … Harrison, J. (2010). Medication reconciliation during internal hospital transfer and impact of computerized prescriber order entry. *Annals of Pharmacotherapy*, *44*(12), 1887–1895.

MacDonald, D., & Neville, D. (2010). Evaluating the implementation of picture archiving and communication systems in Newfoundland and Labrador: A cost benefit analysis. *Journal of Digital Imaging*, *23*(6), 721–731.

McGinn, C. A., Gagnon, M., Shaw, N., Sicotte, C., Mathieu, L., Leduc, Y., … Legare, F. (2012). Users' perspectives of key factors to implementing electronic health records in Canada: a delphi study. *BMC Medical Informatics and Decision Making*, *12*, 105.

Mensink, N., & Paterson, G. (2010). The evolution and uptake of drug information system: The case of a small Canadian province. *Studies in Health Technology & Informatics*, *160*(Pt. 1), 352–355.

Office of the Auditor General of Canada [OAG]. (2010, April). *Electronic health records in Canada: An overview of federal and provincial audit reports*. Ottawa: Author. Retrieved from http://www.oag-bvg.gc.ca/internet/docs/parl_oag_201004_07_e.pdf

O'Reilly, D., Holbrook, A., Blackhouse, G., Troyan, S., & Goeree, R. (2012). Cost-effectiveness of a shared computerized decision support system for diabetes linked to electronic medical records. *Journal of the American Medical Informatics Association*, *19*(3), 341–345.

Paré, G., de Guinea, A. O., Raymond, L., Poba-Nzaou, P., Trudel, M., Marsan, J., & Micheneau, T. (2013). *Computerization of primary care medical clinics in Quebec: Results from a survey on EMR adoption, use and impacts*. Toronto: Canada Health Infoway. Retrieved from https://www.infoway-inforoute.ca/index.php/resources/reports

Paterson, G., Shaw, N., Grant, A., Leonard, K., Delisle, E., Mitchell, S., … Kraetschmer, N. (2010). A conceptual framework for analyzing how Canadian physicians are using electronic medical records in clinical care. *Studies in Health Technology & Informatics*, *160*(Pt. 1), 141–145.

Physician Information Technology Office [PITO]. (2013). *EMR adoption study*. Vancouver: Insights West and Cientis Technologies.Retrieved from https://www.infoway-inforoute.ca/index.php/resources/reports

Poissant, L., Taylor, L., Huang, A., & Tamblyn, R. (2010). Assessing the accuracy of an inter-institutional automated patient-specific health problem list. *BMC Medical Informatics and Decision Making*, *10*, 10.

Prairie Research Associates [PRA]. (2012). *Manitoba's physician integrated network (PIN) initiative: A benefits evaluation report*. Winnipeg, MB: Author. Retrieved from https://www.infoway-inforoute.ca/index.php/resources/reports

PricewaterhouseCoopers [PWC]. (2013). *The emerging benefits of electronic medical record use in community-based care.* Toronto: Canada Health Infoway. Retrieved from https://www.infoway-inforoute.ca/index.php/resources/reports

Price, M., Bowen, M., Lau, F., Kitson, N., & Bardal, S. (2012). Assessing accuracy of an electronic provincial medication repository. *BMC Medical Informatics and Decision Making*, *12*, 42.

Prince Edward Island (P.E.I.) Department of Health and Wellness. (2010). *Prince Edward Island drug information system evaluation report.* Charlottetown: Author. Retrieved from https://www.infoway-inforoute.ca/index.php/resources/reports

Roshanov, P., Gerstein, H. C., Hunt, D. L., Sebaldt, R. J., & Haynes, R. B. (2012). Impact of a computerized system for evidence-based diabetes care on completeness of records: A before-after study. *BMC Medical Informatics and Decision Making*, *12*, 63. doi: 10.1186/1472-6947-12-63

Rozenblum, R., Jang, Y., Zimlichman, E., Salzberg, C., Tamblyn, M., Buckeridge, D., … Tamblyn, R. (2011). A qualitative study of Canada's experience with the implementation of electronic health information. *Canadian Medical Association Journal*, *183*(5), E281–E288.

Schryen, G. (2013). Revisiting IS business value research: what we already know, what we still need to know, and how we can get there. *European Journal of Information Systems*, *22*(2), 139–169.

Shachak, A., Montgomery, C., Tu, K., Jadad, A. R., & Lemieux-Charles, L. (2013). End-user support for a primary care electronic medical record: a qualitative case study of a vendor's perspective. *Informatics in Primary Care*, *20*(3), 185–196.

Sicotte, C., & Paré, G. (2010). Success in health information exchange projects: solving the implementation puzzle. *Social Science & Medicine*, *70*(8), 1159–1165.

Syed, S., Wang, D., Goulard, D., Rich, T., Innes, G., & Lang, E. (2013). Computer order entry systems in the emergency department significantly reduce the time to medication delivery for high acuity patients. *International Journal of Emergency Medicine*, *6*, 20. doi: 10.1186/1865-1380-6-20

Tamblyn, R., Reidel, K., Huang, A., Taylor, L., Winslade, N., Bartlett, G., … Pinsonneault, A. (2010). Increasing the detection and response to adherence problems with cardiovascular medication in primary care through computerized drug management systems: A randomized control trial. *Medical Decision Making*, *30*(2), 176–188.

Terry, A. L., Brown, J. B., Denomme, L. B., Thind, A., & Stewart, M. (2012). Perspectives on electronic medical record implementation after two years of use in primary health care practice. *The Journal of the American Board of Family Medicine*, *25*(4), 522–527.

Tolar, M., & Balka, E. (2011). Beyond individual patient care: enhanced use of EMR data in a primary care setting. *Studies in Health Technology & Informatics*, *164*, 143–147.

Tu, K., Mitiku, T., Lee, D. S., Guo, H., & Tu, J. V. (2010a). Validation of physician billing and hospitalization data to identify patients with ischemic heart disease using data from the electronic medical record administrative data linked database (EMRALD). *Canadian Journal of Cardiology*, *26*(7), e225–e228.

Tu, K., Mitiku, T., Guo, H., Lee, D. S., & Tu, J. V. (2010b). Myocardial infarction and the validation of physician billing and hospitalization data using electronic medical records. *Chronic Diseases in Canada*, *30*(4), 141–146.

Tu, K., Manuel, D., Lam, K., Kavanaugh, D., Mitiku, T. F., & Guo, H. (2011). Diabetics can be identified in an electronic medical record using laboratory tests and prescriptions. *Journal of Clinical Epidemiology*, *64*(4), 431–435.

Part II

Methodological Details

Chapter 8

Methodological Landscape for eHealth Evaluation

Craig Kuziemsky, Francis Lau

8.1 Introduction

This chapter provides the methodological landscape for eHealth evaluation. We will introduce the philosophical assumptions and approaches by which eHealth evaluation is based. We will describe different evaluation methods that have been reported in the literature. We will also include good practices and reporting guidance as ways to advance the field of eHealth evaluation.

Evaluation, broadly defined, needs to answer this question: How do we know this works? A starting point for the conceptual foundation of eHealth evaluation is to ask ourselves what we want to evaluate. Domains outside of healthcare (e.g., manufacturing, retail, finance) are often used as points of comparison for the design, implementation and evaluation of health information technology (HIT). However, a key difference is that in other domains, the goal of IT is typically to automate a specific process and to determine how well the process of automation works as the basis for evaluation. For example, UPS used IT to develop models for predictive analytics to maintain its truck fleet, while Wal-Mart developed a sophisticated supply chain system to link retail and supply elements of its business (Nash, 2015). In these examples, evaluating the IT implementation is relatively straightforward as the objective is to evaluate a process that is both mature and defined.

Evaluating eHealth systems is far more challenging for several reasons. Foremost is that we often do not measure a single process. Rather, healthcare processes are often multifaceted and complex and evaluation must understand and incorporate that complexity into the evaluation process (Kannampallil, Schauer, Cohen, & Patel, 2011). One example is a process like collaboration,

which is a common objective of eHealth systems, but consists of many subprocesses (Eikey, Reddy, & Kuziemsky, 2015). Another challenge is that many of the processes we are trying to support through eHealth may be lacking maturity and thus we need to account for a time component when we design an evaluation strategy.

8.2 Philosophical Assumptions and Approaches

8.2.1 Objectivist and Subjectivist Traditions

Within evaluation research, two predominant philosophical traditions exist: The objectivist and the subjectivist traditions (Friedman & Wyatt, 2014). The objectivist tradition comes from the positivist paradigm, also referred to as "logical science", and assumes that reality is objectively given and can be described by measurable properties, which are independent of the observer (researcher) and his or her instruments (Creswell, 2013). The subjectivist paradigm posits that reality cannot always be measured precisely but rather depends on the observer. It is possible that different observers may have different opinions about the impact and outcome of an implementation (Friedman & Wyatt, 2014).

Early eHealth evaluation was largely influenced by the randomized controlled trial (RCT) research design that predominated in medicine for the evaluation of drugs and therapies. HIT was thought to be another intervention that could be measured and evaluated using controlled conditions to isolate a particular intervention. However, over time it was shown that these approaches do not work well for evaluating the complex multifaceted nature of eHealth implementation (Kaplan, 2001; Koppel, 2015). The controlled RCT environment may not be suitable for evaluating the complex and messy reality where eHealth systems are used. Seminal work by Ash and colleagues identified how Computer Physician Order Entry (CPOE) implementation may lead to unintended consequences (Ash et al., 2003; Ash et al., 2007). While it could be argued that the CPOE system they evaluated was successful from an objective perspective, in that it facilitated automation of orders, it also led to a host of other issues beyond order entry itself, such as communication and workflow issues, changes in the power structure, and the creation of new work. These unintended consequences emphasized the fact that the evaluation of HIT must go beyond just the objective process being automated to also consider the contextual environment where HIT is used (Harrison, Koppel, & Bar-Lev, 2007).

8.2.2 Quantitative versus Qualitative Methods

The evaluation of eHealth systems has spanned the entire spectrum of methodologies and approaches including qualitative, quantitative and mixed methods approaches. Quantitative approaches are useful when we want to evaluate specific aspects of an information system that are independent, objective, and discrete entities (Kaplan & Maxwell, 2005). Examples of variables that can be

measured by quantitative methods include: to study costs and/or benefits; the time taken to complete a task; and the number of patient assessments conducted over a given period (Kaplan, 2001; Kaplan & Maxwell, 2005). Quantitative methods provide an understanding of what has happened.

However, as described above, even if an eHealth system has a favourable objective evaluation, it does not necessarily mean the system is a success. We turn to qualitative studies when we want to evaluate the broader context of system use, or determine whether the evaluation should study issues that are not easily reduced into an objective variable (Kaplan & Maxwell, 2005; Friedman & Wyatt, 2014). Qualitative methods allow an evaluation to encompass meaning and context of the system being studied, and the specific events and processes that define how a system is used over time, in real-life natural settings (Maxwell, 2013). Commonly used qualitative approaches include ethnography, which has proven useful for understanding the front-line contexts and circumstances where eHealth systems are used. Overall, qualitative methods are valuable for understanding *why* and *how* things happen.

The relationship between quantitative and qualitative studies is often a source of controversy or debate. Those who favour quantitative approaches may believe that qualitative approaches are "soft" or lack methodological rigour. Those who favour qualitative approaches counter that quantitative approaches provide numbers but not an understanding of the contextual circumstance where a system is used, at times arguing that technologically sound systems may still fail because of user resistance (Koppel, 2015).

In reality, the two methods should be seen as complementary rather than competitive. Mixed method approaches provide a happy medium between quantitative and qualitative approaches. As described above, while quantitative approaches like RCTs are the gold standard for evaluation, they are not practical as an evaluation method on their own because of the need to consider context in HIT evaluation. Similarly, qualitative approaches have shortcomings, most specifically a lack of generalizability and an inability to know the frequency by which criteria occur. Mixed methods provide a way of leveraging the strengths of qualitative and quantitative approaches while mitigating the weaknesses in both methods. Qualitative approaches can provide an initial evaluation of a system and allow the construction of models based on the evaluation. These models then serve as theories that can be tested using quantitative approaches. An example of mixed methods research in eHealth evaluation is the aforementioned CPOE research by Ash and colleagues (Ash et al., 2003; Ash et al., 2007). They first used qualitative approaches to identify and understand significant unintended consequences of CPOE implementation, and then turned to quantitative approaches both to determine frequencies of the unintended consequences and to compare frequencies across different settings.

While mixed method approaches can be a useful approach for eHealth evaluation they can be methodologically challenging. Mixed methods research does not merely involve researchers taking miscellaneous parts from quantitative

and qualitative approaches; rather they must ensure that such studies are done with the necessary rigour (Carayon et al., 2015). Therefore, there is need to ensure that studies draw upon the formal literature on mixed methods research to further expand the evidence base on mixed methods studies.

8.2.3 Formative and Summative Evaluation

HIT implementation has been described as a journey rather than a destination (McDonald, Overhage, Mamlin, Dexter, & Tierney, 2004). In that context, eHealth evaluation must have formative and summative components that evaluate the how a system is used over time. While summative evaluation is necessary to determine whether a system has met its ultimate objectives, it is also necessary to conduct formative evaluation at various points during a system implementation. Many of the processes that we are trying to evaluate — such as collaborative care delivery or patient-centred care — are in an immature or developmental state and thus eHealth tools may need to be designed and evaluated in stages as these processes mature and evolve (Eikey et al., 2015). Another reason is that while users may initially adopt HIT features in a limited way, the repertoire of how they use a system expands over time. One study showed how after implementation an EHR system was used mainly as a documentation tool despite being designed to support organizational goals of care coordination (Sherer, Meyerhoefer, Sheinberg, & Levick, 2015). However, over time as the system was adapted, users began to expand the functionality of its use to include coordination activities. Had the EHR system been evaluated early in its implementation it likely would have yielded unsatisfactory results because of the limited manner it was being used, highlighting the need for ongoing formative evaluation.

Part of formative evaluation is also evaluating the impact that HIT has on processes that are supplementary to the process being automated. While studies of specific technologies and processes (e.g., EHR and/or CPOE systems and data entry) are important, it is equally important that we evaluate the supplementary processes (e.g., communication) of order entry or decision support. While patient safety and collaboration are common objectives for healthcare delivery, Wu and colleagues state how studies of CPOE far outnumber studies of communication and communication technologies, even though communication is a much more prevalent process (Wu et al., 2014). Further, inadequate communication has been shown to impair CPOE processes (Ash et al., 2003), and thus it should be seen as a formative component of CPOE evaluation.

8.2.4 eHealth System Life Cycles

Formative evaluation is easier to do if there is a framework to provide grounding for how and/or when it should be done. One such framework is the System Development Life Cycle (SDLC) that defines system development according to the following phases: planning, analysis, design, implementation and support, and maintenance. In the traditional SDLC all the above phases would be done in linear fashion with most of the evaluation occurring in the final stages of the

cycle. However, this approach was shown to be problematic for HIT design because of the complexity and dynamic nature of system requirements in health care (Kushniruk, 2002). To address that issue, we have seen the development of number of system design approaches that use evaluation methods *throughout* the SDLC. The advantage of that approach is it incorporates continuous formative evaluation to enable redesign should system requirements change.

One example of applying evaluation methods throughout the SDLC is provided by Kushniruk and Patel (2004) where they use the SDLC to frame when different types of usability testing should be done, ranging from exploratory tests at the needs analysis phase, assessment of prototypes at the system design phase, and finally to validation testing at the maintenance phase. Explicitly mapping evaluation methods to the different phases of the SDLC help ensure that formative evolution is thorough and complete.

A similar approach is offered by Friedman and Wyatt (2014) who developed a typology of evaluation approaches that range from evaluating the need for the resource or tool being developed, to the design and usability of the tool, and finally to evaluating the impact of the resource or tool itself. Friedman and Wyatt then supplement the evaluation approaches with a generic structure to be used for evaluation studies. The structure starts with negotiating the aim and objective of an evaluation structure and then proceeds to develop a study design to measure the objectives.

8.3 Types of Evaluation Methods

Evaluation methods can be broadly classified into methods that were developed specifically for different types of HIT and more general evaluation methods

8.3.1 Evaluation Methods Specific for HIT

A number of evaluation frameworks have been developed specific for HIT. For example, Lau, Hagens, and Muttitt (2007) developed the Infoway Benefits Evaluation Framework discussed in detail in chapter 2. This framework is based on the DeLone and McLean information systems success model and includes three dimensions of quality (system, information and service), two dimensions of system usage (use and user satisfaction), and three dimensions of net benefits (quality, access and productivity). Given the requirement for and emphasis on understanding how eHealth systems impact users at point of care, a significant methodological breakthrough in eHealth evaluation was the incorporation of approaches from usability engineering, an example being usability testing, into the design of HIT (Kushniruk & Patel, 2004). These approaches have been beneficial for identifying how eHealth systems impact users during specific tasks (e.g., data entry, medication order) and how usability issues can lead to medical errors and other patient safety issues (Kushniruk, Triola, Borycki, Stein, & Kannry, 2005).

Aside from the evaluation frameworks described above which identify specific aspects of designing or implementing HIT, there are also HIT evaluation frameworks that provide broader considerations for evaluation. Examples of such frameworks include a classification framework by Currie (2005) that identified four general categories that evaluation can be based on: behavioural; social; software development life cycle; and none of the above.

8.3.2 Other Evaluation Methods used in HIT

Aside from evaluation approaches from within the medical informatics community, there are also supplementary communities that have contributed substantially to eHealth evaluation. Examples include fields on the periphery of medical informatics such as Computer Supported Cooperative Work (CSCW) and Human Computer Interaction (HCI). Approaches from CSCW or HCI are popular in HIT studies as they both focus on the manner and contexts in which people, processes and technology interact, which is a key consideration in eHealth evaluation studies (Pratt, Reddy, McDonald, Tarczy-Hornoch, & Gennari, 2004; Fitzpatrick & Ellingsen, 2013).

Evaluation frameworks adopted and adapted from the management information systems discipline are also popular in HIT studies (Yusof, Papazafeiropoulou, Paul, & Stergioulas, 2007). Examples of such frameworks include Activity Theory, Actor Network Theory, and the Delone and McLean information systems success model (Sadgehi, Andreev, Benyoucef, Momtahan, & Kuziemsky, 2014; Bossen, Jensen, & Udsen, 2013).

The chapters in this section of the handbook — subtitled "Methodological Details" — all describe different evaluation approaches that are relevant to HIT studies. In chapter 8, Kuziemsky and Lau set the stage for this section by introducing the methodological landscape for eHealth evaluation. In chapter 9, Paré and Kitsiou describe approaches for conducting literature reviews for the evaluation of scientific literature. In chapter 10, Lau and Holbrook present methods for conducting comparative studies. In chapter 11, Gu and Warren discuss how descriptive studies contribute to the evaluation of eHealth systems in terms of the system planning, design, implementation, use and impact. In chapter 12, Lau outlines how correlational studies can enhance eHealth evaluation. In chapter 13, Lau discusses methods for survey studies, while in chapter 14 he outlines the economic evaluation of HIT and how to determine whether an eHealth investment provides "value for money". In chapter 15, Anderson and Fu introduce modelling and simulation methods and the role they can play in eHealth studies. In chapter 16, Lau returns to describe approaches to eHealth data quality assessment that are relevant to healthcare organizations. Finally, in chapter 17, Kuziemsky and Lau summarize the key messages from this section, discuss the complexity of HIT implementation, and offer insight as to good eHealth evaluation practices in light of the complexity. Taken together, these chapters provide a diverse perspective on eHealth evaluation that spans the entire SDLC from literature retrieval and simulation as part of deriving requirements to descriptive

and comparative studies of eHealth implementation to evaluating the economic impact and data quality of eHealth systems.

8.4 Methodological Guidance

While eHealth evaluation has benefited from the breadth of evaluation methods as discussed in the previous section, one of the challenges with a broad base of evaluation methods is that a lack of consistency or quality standardization prevents the sharing of evaluation outcomes across different settings (Brender et al., 2013). This lack of comparability can be important in that it may prevent meaningful comparison of potentially significant findings. For example, two studies that identified contradictory findings about CPOE usage could not be compared and the discrepancies reconciled because of significant differences in the HIT evaluation research design (Ammenwerth et al., 2006).

To address methodological issues around evaluation and comparability, frameworks have been developed to provide consistency and quality in the reporting of HIT evaluation results. One such framework, the STARE-HI statement, was developed to enhance how qualitative and quantitative evaluation studies are reported (Brender et al., 2013). Following the STARE-HI principles enables studies to be contextually evaluated to permit readers of such papers to better place the studies in a proper context and judge their validity and generalizability. The STARE-HI statement specifies which items should be contained in a publication of a health informatics evaluation study in order to enable others to judge the trustworthiness of a study's establishment, its design, its execution and line of reasoning, and the validity of its conclusion, as well as its context and thus the potential for generalizability.

Another framework is the guideline for good evaluation practice in health informatics (GEP-HI) developed by Nykänen and colleagues (2011). GEP-HI consists of a list of 60 items that are relevant for planning, implementation and execution of an eHealth evaluation study. The items include budgets, ethical and legal considerations, identification and recruitment of participants, risk management and project control and the undertaking of the evaluation study and reporting of results. To aid in the practicality of the application of these items, they are framed around the different phases of an evaluation study: preliminary outline, study design, operationalization of methods, project planning, and execution and completion of the evaluation study (Nykänen et al., 2011).

8.5 Issues and Challenges

This chapter has introduced the eHealth evaluation landscape and some of the methods and frameworks used in eHealth evaluation. While there are many potential evaluation approaches, a significant challenge is determining what approach to use for a particular evaluation study. The first step in determining the right approach is identifying what an evaluation study needs to report on. For

example, an economic evaluation, using the methods described in chapter 13 can evaluate the economic return on a system, but will provide no insight into how the system interacts with users or care processes. Similarly, if an evaluation study looks at how a system has impacted process efficiency, it is possible that a process (e.g., order entry or patient discharge) may become more efficient via automation (and thus would have a favourable evaluation outcome) but still cause workflow or communication issues.

The bottom line is a study cannot evaluate all possible outcomes and it is important to be very clear on the question of "what". In eHealth evaluation, therefore, the first question that must be asked is: What do we want to evaluate? This question is often not straightforward. Patient safety and collaborative care delivery are desired objectives of healthcare delivery and thus the eHealth systems we design should support care delivery, safety and collaboration; but these are also abstract concepts and not easily measurable. We cannot consider patient safety per se, as safety is comprised of multiple factors. Rather, we need to define the underlying processes that measure the underlying processes that influence patient safety.

We summarize the issues and challenges presented in this chapter as three considerations for eHealth evaluation. First is the need to understand the complexity of the healthcare system. Healthcare can be classified as a complex adaptive system (CAS) because the various elements within it — such as care delivery, education, and policy — consist of a series of interacting parts that work in nonlinear and evolving ways (Kannampallil et al., 2011; Kuziemsky, 2016). A challenge with a CAS is that it is not always possible to predict how different parts will interact in a given situation. Introducing automation for a particular process may have profound or unexpected impacts on other processes that could not be anticipated ahead of time. The more system components an HIT may interact with, the more wide reaching the evaluation needs to be. Multilevel evaluation studies are often necessary to understand the impact that an eHealth system may have at such different levels as those of the individual provider and the healthcare team.

The second consideration is defining what method is best suited to achieve our evaluation objectives. A common debate in eHealth evaluation is whether a qualitative or quantitative approach should be used. However, we suggest that such arguments are not helpful and, rather, the approaches should be looked as complementary to each other. As described earlier, both approaches have strengths and weaknesses and the key is to leverage the strengths of both approaches. If we are doing an exploratory study (such as assessing how eHealth implementation impacts a clinical unit) then qualitative methods are better suited as they enable us to gain an understanding of what is occurring and why it occurs. However, again as stated earlier, mixed methods approaches should be used to then quantify the significance of the impacts.

The third consideration is the need to understand that eHealth evaluation is almost always time limited because of the evolving nature of healthcare pro-

cesses and technologies. As described earlier, domains such as manufacturing and retail have succeeded at IT-enabled automation largely because they are automating well-structured and well-defined processes; eHealth is typically automating immature processes (e.g., collaboration) and thus a multi-time evaluation may be needed in order to evaluate a process over time.

8.6 Summary

Evaluating eHealth systems is challenging because of the complexity of healthcare delivery. However, there is a wide body of research and evidence to guide eHealth evaluation. This chapter outlined the philosophical assumptions and approaches and specific evaluation methods for evaluating eHealth systems, as well as providing methodological guidance for carrying out eHealth evaluations.

References

Ammenwerth, E., Talmon, J., Ash, J. S., Bates, D. W., Beuscart-Zéphir, M. C., Duhamel, A., … Geissbuhler, A. (2006). Impact of CPOE on mortality rates—contradictory findings, important messages. *Methods of Information in Medicine, 45*(6), 586–593.

Ash, J. S., Gorman, P. N., Lavelle, M., Stavri, P. Z., Lyman, J., Fournier, L., & Carpenter, J. (2003). Perceptions of physician order entry: Results of a cross-site qualitative study. *Methods of Information in Medicine, 42*(4), 313–323.

Ash, J. S., Sittig, D. F., Poon, E. G., Guappone, K., Campbell, E., & Dykstra, R. H. (2007). The extent and importance of unintended consequences related to computerized provider order entry. *Journal of the American Medical Informatics Association, 4*(4), 415–423.

Bossen, C., Jensen, L. G., & Udsen, F. W. (2013). Evaluation of a comprehensive EHR based on the DeLone and McLean model for IS success: Approach, results, and success factors. *International Journal of Medical Informatics, 82*(10), 940–953.

Carayon, P., Kianfar, S., Li, Y., Xie, A., Alyousef, B., & Wooldridge, A. (2015). A systematic review of mixed methods research on human factors and ergonomics in health care. *Applied Ergonomics, 51*, 291–321.

Brender, J., Talmon, J., de Keizer, N., Nykänen, P., Rigby, M., & Ammenwerth, E. (2013). STARE-HI: Statement on reporting of evaluation studies in health informatics, explanation and elaboration. *Applied Clinical Informatics, 4*(3), 331–358.

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. Thousand Oaks, CA: SAGE Publications.

Currie, L. M. (2005). Evaluation frameworks for nursing informatics. *International Journal of Medical Informatics, 74*(11–12 Nursing Informatics Special Issue), 908–916.

Eikey, E. V., Reddy, M. C., & Kuziemsky, C. E. (2015). Examining the role of collaboration in studies of health information technologies in biomedical informatics: A systematic review of 25 years of research. *Journal of Biomedical Informatics, 57*, 263–277.

Fitzpatrick, G., & Ellingsen, G. (2013). A review of 25 years of CSCW research in healthcare: contributions, challenges and future agendas. *Computer Supported Cooperative Work, 22*(4/6), 609–665.

Friedman, C. P., & Wyatt, J. C. (2014). Evaluation of biomedical and health information resources. In E. H. Shortliffe & J. J. Cimino (Eds.), *Biomedical informatics* (pp. 355–387). London: Springer-Verlag.

Harrison, M. I., Koppel, R., & Bar-Lev, S. (2007). Unintended consequences of information technologies in health care: An interactive sociotechnical analysis. *Journal of the American Medical Informatics Association, 14*(5), 542–549.

Kannampallil, T., Schauer, G., Cohen, T., & Patel, V. (2011). Considering complexity in healthcare systems. *Journal of Biomedical Informatics, 44*(6), 943–947.

Kaplan, B. (2001). Evaluating informatics applications — some alternative approaches: theory, social interactionism and call for methodological pluralism. *International Journal of Medical Informatics, 64*(1), 39–56.

Kaplan, B., & Maxwell, J. A. (2005). Qualitative research methods for evaluating computer information systems. In J. G. Anderson, C. E. Aydin, & S. J. Jay (Eds.), *Evaluating the organizational impact of healthcare information systems* (2nd ed., pp. 30–55). Newbury Park, CA: SAGE Publications.

Koppel, R. (2015). Great promises of healthcare information technology deliver less. In C. A. Weaver, M. J. Ball, G. R. Kim, & J. M. Kiel (Eds.), *Healthcare information management systems: Cases, strategies, and solutions* (pp. 101–125). New York: Springer.

Kushniruk, A. (2002). Evaluation in the design of health information systems: Application of approaches emerging from usability engineering. *Computers in Biology and Medicine, 32*(3), 141–149.

Kushniruk, A., & Patel, V. (2004). Cognitive and usability engineering methods for the evaluation of clinical information systems. *Journal of Biomedical Informatics, 37*(1), 56–76.

Kushniruk, A. W., Triola, M. M., Borycki, E. M., Stein, B., & Kannry, J. L. (2005). Technology induced error and usability: the relationship between usability problems and prescription errors when using a handheld application. *International Journal of Medical Informatics, 74*(7/8), 519–526.

Kuziemsky, C. E. (2016). Decision-making in healthcare as a complex adaptive system. *Healthcare Management Forum, 29*(1), 4–7.

Lau, F., Hagens, S., & Muttitt, S. (2007). A proposed benefits evaluation framework for health information systems in Canada. *Healthcare Quarterly, 10*(1), 112–118.

McDonald, C. J., Overhage, J. M., Mamlin, B. W., Dexter, P. D., & Tierney, W. M. (2004). Physicians, information technology, and health care systems: A journey, not a destination. *Journal of the American Medical Informatics Association, 11*(2), 121–124.

Maxwell, J. A. (2013). *Qualitative research design: An interactive approach* (3rd ed.). Thousand Oaks, CA: SAGE Publications.

Nash, K. (2015, May 7). Wal-Mart builds supply chain to meet e-commerce demands. *Wall Street Journal*. Retrieved from http://www.wsj.com/articles/wal-mart-builds-supply-chain-to-meet-e-commerce-demands-1431016708

Nykänen, P., Brender, J., Talmon, J., de Keizer, N., Rigby, M., Beuscart-Zephir, M. C., & Ammenwerth, E. (2011). Guideline for good evaluation practice in health informatics (GEP-HI). *International Journal of Medical Informatics, 80*(12), 815–827.

Pratt, W., Reddy, M. C., McDonald, D. W., Tarczy-Hornoch, P., & Gennari, J. H. (2004). Incorporating ideas from computer-supported cooperative work. *Journal of Biomedical Informatics, 37*(2), 128–137.

Sadeghi, P., Andreev, P., Benyoucef, M., Momtahan, K., & Kuziemsky, C. (2014). Activity theory driven system analysis of complex healthcare processes. In *Proceedings of the 22nd European Conference on Information Systems (ECIS)* (pp. 1–14), June 9 to 11, Tel Aviv, Israel.

Sherer, S. A., Meyerhoefer, C. D., Sheinberg, M., & Levick, D. (2015). Integrating commercial ambulatory electronic health records with hospital systems: An evolutionary process. *International Journal of Medical Informatics, 84*(9), 683–693.

Wu, R., Appel, L., Morra, D., Lo, V., Kitto, S., & Quan, S. (2014). Short message service or disService: Issues with text messaging in a complex medical environment. *International Journal of Medical Informatics, 83*(4), 278–284.

Yusof, M. M., Papazafeiropoulou, A., Paul, R. J., & Stergioulas, L. K. (2008). Investigating evaluation frameworks for health information systems. *International Journal of Medical Informatics, 77*(6), 377–385.

Chapter 9

Methods for Literature Reviews

Guy Paré, Spyros Kitsiou

9.1 Introduction

Literature reviews play a critical role in scholarship because science remains, first and foremost, a cumulative endeavour (vom Brocke et al., 2009). As in any academic discipline, rigorous knowledge syntheses are becoming indispensable in keeping up with an exponentially growing eHealth literature, assisting practitioners, academics, and graduate students in finding, evaluating, and synthesizing the contents of many empirical and conceptual papers. Among other methods, literature reviews are essential for: (a) identifying what has been written on a subject or topic; (b) determining the extent to which a specific research area reveals any interpretable trends or patterns; (c) aggregating empirical findings related to a narrow research question to support evidence-based practice; (d) generating new frameworks and theories; and (e) identifying topics or questions requiring more investigation (Paré, Trudel, Jaana, & Kitsiou, 2015).

Literature reviews can take two major forms. The most prevalent one is the "literature review" or "background" section within a journal paper or a chapter in a graduate thesis. This section synthesizes the extant literature and usually identifies the gaps in knowledge that the empirical study addresses (Sylvester, Tate, & Johnstone, 2013). It may also provide a theoretical foundation for the proposed study, substantiate the presence of the research problem, justify the research as one that contributes something new to the cumulated knowledge, or validate the methods and approaches for the proposed study (Hart, 1998; Levy & Ellis, 2006).

The second form of literature review, which is the focus of this chapter, constitutes an original and valuable work of research in and of itself (Paré et al., 2015). Rather than providing a base for a researcher's own work, it creates a solid starting point for all members of the community interested in a particular area

or topic (Mulrow, 1987). The so-called "review article" is a journal-length paper which has an overarching purpose to synthesize the literature in a field, without collecting or analyzing any primary data (Green, Johnson, & Adams, 2006).

When appropriately conducted, review articles represent powerful information sources for practitioners looking for state-of-the art evidence to guide their decision-making and work practices (Paré et al., 2015). Further, high-quality reviews become frequently cited pieces of work which researchers seek out as a first clear outline of the literature when undertaking empirical studies (Cooper, 1988; Rowe, 2014). Scholars who track and gauge the impact of articles have found that review papers are cited and downloaded more often than any other type of published article (Cronin, Ryan, & Coughlan, 2008; Montori, Wilczynski, Morgan, Haynes, & Hedges, 2003; Patsopoulos, Analatos, & Ioannidis, 2005). The reason for their popularity may be the fact that reading the review enables one to have an overview, if not a detailed knowledge of the area in question, as well as references to the most useful primary sources (Cronin et al., 2008). Although they are not easy to conduct, the commitment to complete a review article provides a tremendous service to one's academic community (Paré et al., 2015; Petticrew & Roberts, 2006). Most, if not all, peer-reviewed journals in the fields of medical informatics publish review articles of some type.

The main objectives of this chapter are fourfold: (a) to provide an overview of the major steps and activities involved in conducting a stand-alone literature review; (b) to describe and contrast the different types of review articles that can contribute to the eHealth knowledge base; (c) to illustrate each review type with one or two examples from the eHealth literature; and (d) to provide a series of recommendations for prospective authors of review articles in this domain.

9.2 Overview of the Literature Review Process and Steps

As explained in Templier and Paré (2015), there are six generic steps involved in conducting a review article:

1. formulating the research question(s) and objective(s),
2. searching the extant literature,
3. screening for inclusion,
4. assessing the quality of primary studies,
5. extracting data, and
6. analyzing data.

Although these steps are presented here in sequential order, one must keep in mind that the review process can be iterative and that many activities can be initiated during the planning stage and later refined during subsequent phases (Finfgeld-Connett & Johnson, 2013; Kitchenham & Charters, 2007).

Formulating the research question(s) and objective(s): As a first step, members of the review team must appropriately justify the need for the review itself (Petticrew & Roberts, 2006), identify the review's main objective(s) (Okoli & Schabram, 2010), and define the concepts or variables at the heart of their synthesis (Cooper & Hedges, 2009; Webster & Watson, 2002). Importantly, they also need to articulate the research question(s) they propose to investigate (Kitchenham & Charters, 2007). In this regard, we concur with Jesson, Matheson, and Lacey (2011) that clearly articulated research questions are key ingredients that guide the entire review methodology; they underscore the type of information that is needed, inform the search for and selection of relevant literature, and guide or orient the subsequent analysis.

Searching the extant literature: The next step consists of searching the literature and making decisions about the suitability of material to be considered in the review (Cooper, 1988). There exist three main coverage strategies. First, exhaustive coverage means an effort is made to be as comprehensive as possible in order to ensure that all relevant studies, published and unpublished, are included in the review and, thus, conclusions are based on this all-inclusive knowledge base. The second type of coverage consists of presenting materials that are representative of most other works in a given field or area. Often authors who adopt this strategy will search for relevant articles in a small number of top-tier journals in a field (Paré et al., 2015). In the third strategy, the review team concentrates on prior works that have been central or pivotal to a particular topic. This may include empirical studies or conceptual papers that initiated a line of investigation, changed how problems or questions were framed, introduced new methods or concepts, or engendered important debate (Cooper, 1988).

Screening for inclusion: The following step consists of evaluating the applicability of the material identified in the preceding step (Levy & Ellis, 2006; vom Brocke et al., 2009). Once a group of potential studies has been identified, members of the review team must screen them to determine their relevance (Petticrew & Roberts, 2006). A set of predetermined rules provides a basis for including or excluding certain studies. This exercise requires a sig-

nificant investment on the part of researchers, who must ensure enhanced objectivity and avoid biases or mistakes. As discussed later in this chapter, for certain types of reviews there must be at least two independent reviewers involved in the screening process and a procedure to resolve disagreements must also be in place (Liberati et al., 2009; Shea et al., 2009).

Assessing the quality of primary studies: In addition to screening material for inclusion, members of the review team may need to assess the scientific quality of the selected studies, that is, appraise the rigour of the research design and methods. Such formal assessment, which is usually conducted independently by at least two coders, helps members of the review team refine which studies to include in the final sample, determine whether or not the differences in quality may affect their conclusions, or guide how they analyze the data and interpret the findings (Petticrew & Roberts, 2006). Ascribing quality scores to each primary study or considering through domain-based evaluations which study components have or have not been designed and executed appropriately makes it possible to reflect on the extent to which the selected study addresses possible biases and maximizes validity (Shea et al., 2009).

Extracting data: The following step involves gathering or extracting applicable information from each primary study included in the sample and deciding what is relevant to the problem of interest (Cooper & Hedges, 2009). Indeed, the type of data that should be recorded mainly depends on the initial research questions (Okoli & Schabram, 2010). However, important information may also be gathered about how, when, where and by whom the primary study was conducted, the research design and methods, or qualitative/quantitative results (Cooper & Hedges, 2009).

Analyzing and synthesizing data: As a final step, members of the review team must collate, summarize, aggregate, organize, and compare the evidence extracted from the included studies. The extracted data must be presented in a meaningful way that suggests a new contribution to the extant literature (Jesson et al., 2011). Webster and Watson (2002) warn researchers that literature reviews should be much more than lists of papers and should provide a coherent lens to make sense of extant knowledge on a given topic. There exist several methods and techniques for synthesizing quantitative (e.g., frequency analysis, meta-analysis) and qualitative (e.g., grounded theory, narrative analysis, meta-ethnography) evidence

(Dixon-Woods, Agarwal, Jones, Young, & Sutton, 2005; Thomas & Harden, 2008).

9.3 Types of Review Articles and Brief Illustrations

EHealth researchers have at their disposal a number of approaches and methods for making sense out of existing literature, all with the purpose of casting current research findings into historical contexts or explaining contradictions that might exist among a set of primary research studies conducted on a particular topic. Our classification scheme is largely inspired from Paré and colleagues' (2015) typology. Below we present and illustrate those review types that we feel are central to the growth and development of the eHealth domain.

9.3.1 Narrative Reviews

The *narrative review* is the "traditional" way of reviewing the extant literature and is skewed towards a qualitative interpretation of prior knowledge (Sylvester et al., 2013). Put simply, a narrative review attempts to summarize or synthesize what has been written on a particular topic but does not seek generalization or cumulative knowledge from what is reviewed (Davies, 2000; Green et al., 2006). Instead, the review team often undertakes the task of accumulating and synthesizing the literature to demonstrate the value of a particular point of view (Baumeister & Leary, 1997). As such, reviewers may selectively ignore or limit the attention paid to certain studies in order to make a point. In this rather unsystematic approach, the selection of information from primary articles is subjective, lacks explicit criteria for inclusion and can lead to biased interpretations or inferences (Green et al., 2006). There are several narrative reviews in the particular eHealth domain, as in all fields, which follow such an unstructured approach (Silva et al., 2015; Paul et al., 2015).

Despite these criticisms, this type of review can be very useful in gathering together a volume of literature in a specific subject area and synthesizing it. As mentioned above, its primary purpose is to provide the reader with a comprehensive background for understanding current knowledge and highlighting the significance of new research (Cronin et al., 2008). Faculty like to use narrative reviews in the classroom because they are often more up to date than textbooks, provide a single source for students to reference, and expose students to peer-reviewed literature (Green et al., 2006). For researchers, narrative reviews can inspire research ideas by identifying gaps or inconsistencies in a body of knowledge, thus helping researchers to determine research questions or formulate hypotheses. Importantly, narrative reviews can also be used as educational articles to bring practitioners up to date with certain topics of issues (Green et al., 2006).

Recently, there have been several efforts to introduce more rigour in narrative reviews that will elucidate common pitfalls and bring changes into their publication standards. Information systems researchers, among others, have contributed to advancing knowledge on how to structure a "traditional" review.

For instance, Levy and Ellis (2006) proposed a generic framework for conducting such reviews. Their model follows the systematic data processing approach comprised of three steps, namely: (a) literature search and screening; (b) data extraction and analysis; and (c) writing the literature review. They provide detailed and very helpful instructions on how to conduct each step of the review process. As another methodological contribution, vom Brocke et al. (2009) offered a series of guidelines for conducting literature reviews, with a particular focus on how to search and extract the relevant body of knowledge. Last, Bandara, Miskon, and Fielt (2011) proposed a structured, predefined and tool-supported method to identify primary studies within a feasible scope, extract relevant content from identified articles, synthesize and analyze the findings, and effectively write and present the results of the literature review. We highly recommend that prospective authors of narrative reviews consult these useful sources before embarking on their work.

Darlow and Wen (2015) provide a good example of a highly structured narrative review in the eHealth field. These authors synthesized published articles that describe the development process of mobile health (m-health) interventions for patients' cancer care self-management. As in most narrative reviews, the scope of the research questions being investigated is broad: (a) how development of these systems are carried out; (b) which methods are used to investigate these systems; and (c) what conclusions can be drawn as a result of the development of these systems. To provide clear answers to these questions, a literature search was conducted on six electronic databases and *Google Scholar*. The search was performed using several terms and free text words, combining them in an appropriate manner. Four inclusion and three exclusion criteria were utilized during the screening process. Both authors independently reviewed each of the identified articles to determine eligibility and extract study information. A flow diagram shows the number of studies identified, screened, and included or excluded at each stage of study selection. In terms of contributions, this review provides a series of practical recommendations for m-health intervention development.

9.3.2 Descriptive or Mapping Reviews

The primary goal of a *descriptive review* is to determine the extent to which a body of knowledge in a particular research topic reveals any interpretable pattern or trend with respect to pre-existing propositions, theories, methodologies or findings (King & He, 2005; Paré et al., 2015). In contrast with narrative reviews, descriptive reviews follow a systematic and transparent procedure, including searching, screening and classifying studies (Petersen, Vakkalanka, & Kuzniarz, 2015). Indeed, structured search methods are used to form a representative sample of a larger group of published works (Paré et al., 2015). Further, authors of descriptive reviews extract from each study certain characteristics of interest, such as publication year, research methods, data collection techniques, and direction or strength of research outcomes (e.g., positive, negative,

or non-significant) in the form of frequency analysis to produce quantitative results (Sylvester et al., 2013). In essence, each study included in a descriptive review is treated as the unit of analysis and the published literature as a whole provides a database from which the authors attempt to identify any interpretable trends or draw overall conclusions about the merits of existing conceptualizations, propositions, methods or findings (Paré et al., 2015). In doing so, a descriptive review may claim that its findings represent the state of the art in a particular domain (King & He, 2005).

In the fields of health sciences and medical informatics, reviews that focus on examining the range, nature and evolution of a topic area are described by Anderson, Allen, Peckham, and Goodwin (2008) as *mapping reviews*. Like descriptive reviews, the research questions are generic and usually relate to publication patterns and trends. There is no preconceived plan to systematically review all of the literature although this can be done. Instead, researchers often present studies that are representative of most works published in a particular area and they consider a specific time frame to be mapped.

An example of this approach in the eHealth domain is offered by DeShazo, Lavallie, and Wolf (2009). The purpose of this descriptive or mapping review was to characterize publication trends in the medical informatics literature over a 20-year period (1987 to 2006). To achieve this ambitious objective, the authors performed a bibliometric analysis of medical informatics citations indexed in MEDLINE using publication trends, journal frequencies, impact factors, Medical Subject Headings (MeSH) term frequencies, and characteristics of citations. Findings revealed that there were over 77,000 medical informatics articles published during the covered period in numerous journals and that the average annual growth rate was 12%. The MeSH term analysis also suggested a strong interdisciplinary trend. Finally, average impact scores increased over time with two notable growth periods. Overall, patterns in research outputs that seem to characterize the historic trends and current components of the field of medical informatics suggest it may be a maturing discipline (DeShazo et al., 2009).

9.3.3 Scoping Reviews

Scoping reviews attempt to provide an initial indication of the potential size and nature of the extant literature on an emergent topic (Arksey & O'Malley, 2005; Daudt, van Mossel, & Scott, 2013; Levac, Colquhoun, & O'Brien, 2010). A scoping review may be conducted to examine the extent, range and nature of research activities in a particular area, determine the value of undertaking a full systematic review (discussed next), or identify research gaps in the extant literature (Paré et al., 2015). In line with their main objective, scoping reviews usually conclude with the presentation of a detailed research agenda for future works along with potential implications for both practice and research.

Unlike narrative and descriptive reviews, the whole point of scoping the field is to be as comprehensive as possible, including grey literature (Arksey & O'Malley, 2005). Inclusion and exclusion criteria must be established to help

researchers eliminate studies that are not aligned with the research questions. It is also recommended that at least two independent coders review abstracts yielded from the search strategy and then the full articles for study selection (Daudt et al., 2013). The synthesized evidence from content or thematic analysis is relatively easy to present in tabular form (Arksey & O'Malley, 2005; Thomas & Harden, 2008).

One of the most highly cited scoping reviews in the eHealth domain was published by Archer, Fevrier-Thomas, Lokker, McKibbon, and Straus (2011). These authors reviewed the existing literature on personal health record (PHR) systems including design, functionality, implementation, applications, outcomes, and benefits. Seven databases were searched from 1985 to March 2010. Several search terms relating to PHRs were used during this process. Two authors independently screened titles and abstracts to determine inclusion status. A second screen of full-text articles, again by two independent members of the research team, ensured that the studies described PHRs. All in all, 130 articles met the criteria and their data were extracted manually into a database. The authors concluded that although there is a large amount of survey, observational, cohort/panel, and anecdotal evidence of PHR benefits and satisfaction for patients, more research is needed to evaluate the results of PHR implementations. Their in-depth analysis of the literature signalled that there is little solid evidence from randomized controlled trials or other studies through the use of PHRs. Hence, they suggested that more research is needed that addresses the current lack of understanding of optimal functionality and usability of these systems, and how they can play a beneficial role in supporting patient self-management (Archer et al., 2011).

9.3.4 Forms of Aggregative Reviews

Healthcare providers, practitioners, and policy-makers are nowadays overwhelmed with large volumes of information, including research-based evidence from numerous clinical trials and evaluation studies, assessing the effectiveness of health information technologies and interventions (Ammenwerth & de Keizer, 2004; Deshazo et al., 2009). It is unrealistic to expect that all these disparate actors will have the time, skills, and necessary resources to identify the available evidence in the area of their expertise and consider it when making decisions. Systematic reviews that involve the rigorous application of scientific strategies aimed at limiting subjectivity and bias (i.e., systematic and random errors) can respond to this challenge.

Systematic reviews attempt to aggregate, appraise, and synthesize in a single source all empirical evidence that meet a set of previously specified eligibility criteria in order to answer a clearly formulated and often narrow research question on a particular topic of interest to support evidence-based practice (Liberati et al., 2009). They adhere closely to explicit scientific principles (Liberati et al., 2009) and rigorous methodological guidelines (Higgins & Green, 2008) aimed at reducing random and systematic errors that can lead to deviations from the

truth in results or inferences. The use of explicit methods allows systematic reviews to aggregate a large body of research evidence, assess whether effects or relationships are in the same direction and of the same general magnitude, explain possible inconsistencies between study results, and determine the strength of the overall evidence for every outcome of interest based on the quality of included studies and the general consistency among them (Cook, Mulrow, & Haynes, 1997). The main procedures of a systematic review involve:

1. Formulating a review question and developing a search strategy based on explicit inclusion criteria for the identification of eligible studies (usually described in the context of a detailed review protocol).

2. Searching for eligible studies using multiple databases and information sources, including grey literature sources, without any language restrictions.

3. Selecting studies, extracting data, and assessing risk of bias in a duplicate manner using two independent reviewers to avoid random or systematic errors in the process.

4. Analyzing data using quantitative or qualitative methods.

5. Presenting results in summary of findings tables.

6. Interpreting results and drawing conclusions.

Many systematic reviews, but not all, use statistical methods to combine the results of independent studies into a single quantitative estimate or summary effect size. Known as *meta-analyses*, these reviews use specific data extraction and statistical techniques (e.g., network, frequentist, or Bayesian meta-analyses) to calculate from each study by outcome of interest an effect size along with a confidence interval that reflects the degree of uncertainty behind the point estimate of effect (Borenstein, Hedges, Higgins, & Rothstein, 2009; Deeks, Higgins, & Altman, 2008). Subsequently, they use fixed or random-effects analysis models to combine the results of the included studies, assess statistical heterogeneity, and calculate a weighted average of the effect estimates from the different studies, taking into account their sample sizes. The summary effect size is a value that reflects the average magnitude of the intervention effect for a particular outcome of interest or, more generally, the strength of a relationship between two variables across all studies included in the systematic review. By statistically combining data from multiple studies, meta-analyses can create more precise and reliable estimates of intervention effects than those derived

from individual studies alone, when these are examined independently as discrete sources of information.

The review by Gurol-Urganci, de Jongh, Vodopivec-Jamsek, Atun, and Car (2013) on the effects of mobile phone messaging reminders for attendance at healthcare appointments is an illustrative example of a high-quality systematic review with meta-analysis. Missed appointments are a major cause of inefficiency in healthcare delivery with substantial monetary costs to health systems. These authors sought to assess whether mobile phone-based appointment reminders delivered through Short Message Service (SMS) or Multimedia Messaging Service (MMS) are effective in improving rates of patient attendance and reducing overall costs. To this end, they conducted a comprehensive search on multiple databases using highly sensitive search strategies without language or publication-type restrictions to identify all RCTs that are eligible for inclusion. In order to minimize the risk of omitting eligible studies not captured by the original search, they supplemented all electronic searches with manual screening of trial registers and references contained in the included studies. Study selection, data extraction, and risk of bias assessments were performed independently by two coders using standardized methods to ensure consistency and to eliminate potential errors. Findings from eight RCTs involving 6,615 participants were pooled into meta-analyses to calculate the magnitude of effects that mobile text message reminders have on the rate of attendance at healthcare appointments compared to no reminders and phone call reminders.

Meta-analyses are regarded as powerful tools for deriving meaningful conclusions. However, there are situations in which it is neither reasonable nor appropriate to pool studies together using meta-analytic methods simply because there is extensive clinical heterogeneity between the included studies or variation in measurement tools, comparisons, or outcomes of interest. In these cases, systematic reviews can use qualitative synthesis methods such as vote counting, content analysis, classification schemes and tabulations, as an alternative approach to narratively synthesize the results of the independent studies included in the review. This form of review is known as *qualitative systematic review.*

A rigorous example of one such review in the eHealth domain is presented by Mickan, Atherton, Roberts, Heneghan, and Tilson (2014) on the use of handheld computers by healthcare professionals and their impact on access to information and clinical decision-making. In line with the methodological guidelines for systematic reviews, these authors: (a) developed and registered with PROSPERO (www.crd.york.ac.uk/PROSPERO/) an a priori review protocol; (b) conducted comprehensive searches for eligible studies using multiple databases and other supplementary strategies (e.g., forward searches); and (c) subsequently carried out study selection, data extraction, and risk of bias assessments in a duplicate manner to eliminate potential errors in the review process. Heterogeneity between the included studies in terms of reported outcomes and measures precluded the use of meta-analytic methods. To this end, the authors resorted to using narrative analysis and synthesis to describe the effectiveness of handheld

computers on accessing information for clinical knowledge, adherence to safety and clinical quality guidelines, and diagnostic decision-making.

In recent years, the number of systematic reviews in the field of health informatics has increased considerably. Systematic reviews with discordant findings can cause great confusion and make it difficult for decision-makers to interpret the review-level evidence (Moher, 2013). Therefore, there is a growing need for appraisal and synthesis of prior systematic reviews to ensure that decision-making is constantly informed by the best available accumulated evidence. *Umbrella reviews*, also known as overviews of systematic reviews, are tertiary types of evidence synthesis that aim to accomplish this; that is, they aim to compare and contrast findings from multiple systematic reviews and meta-analyses (Becker & Oxman, 2008). Umbrella reviews generally adhere to the same principles and rigorous methodological guidelines used in systematic reviews. However, the unit of analysis in umbrella reviews is the systematic review rather than the primary study (Becker & Oxman, 2008). Unlike systematic reviews that have a narrow focus of inquiry, umbrella reviews focus on broader research topics for which there are several potential interventions (Smith, Devane, Begley, & Clarke, 2011). A recent umbrella review on the effects of home telemonitoring interventions for patients with heart failure critically appraised, compared, and synthesized evidence from 15 systematic reviews to investigate which types of home telemonitoring technologies and forms of interventions are more effective in reducing mortality and hospital admissions (Kitsiou, Paré, & Jaana, 2015).

9.3.5 Realist Reviews

Realist reviews are theory-driven interpretative reviews developed to inform, enhance, or supplement conventional systematic reviews by making sense of heterogeneous evidence about complex interventions applied in diverse contexts in a way that informs policy decision-making (Greenhalgh, Wong, Westhorp, & Pawson, 2011). They originated from criticisms of positivist systematic reviews which centre on their "simplistic" underlying assumptions (Oates, 2011). As explained above, systematic reviews seek to identify causation. Such logic is appropriate for fields like medicine and education where findings of randomized controlled trials can be aggregated to see whether a new treatment or intervention does improve outcomes. However, many argue that it is not possible to establish such direct causal links between interventions and outcomes in fields such as social policy, management, and information systems where for any intervention there is unlikely to be a regular or consistent outcome (Oates, 2011; Pawson, 2006; Rousseau, Manning, & Denyer, 2008).

To circumvent these limitations, Pawson, Greenhalgh, Harvey, and Walshe (2005) have proposed a new approach for synthesizing knowledge that seeks to unpack the mechanism of how "complex interventions" work in particular contexts. The basic research question — what works? — which is usually associated with systematic reviews changes to: what is it about this intervention that works,

for whom, in what circumstances, in what respects and why? Realist reviews have no particular preference for either quantitative or qualitative evidence. As a theory-building approach, a realist review usually starts by articulating likely underlying mechanisms and then scrutinizes available evidence to find out whether and where these mechanisms are applicable (Shepperd et al., 2009). Primary studies found in the extant literature are viewed as case studies which can test and modify the initial theories (Rousseau et al., 2008).

The main objective pursued in the realist review conducted by Otte-Trojel, de Bont, Rundall, and van de Klundert (2014) was to examine how patient portals contribute to health service delivery and patient outcomes. The specific goals were to investigate how outcomes are produced and, most importantly, how variations in outcomes can be explained. The research team started with an exploratory review of background documents and research studies to identify ways in which patient portals may contribute to health service delivery and patient outcomes. The authors identified six main ways which represent "educated guesses" to be tested against the data in the evaluation studies. These studies were identified through a formal and systematic search in four databases between 2003 and 2013. Two members of the research team selected the articles using a pre-established list of inclusion and exclusion criteria and following a two-step procedure. The authors then extracted data from the selected articles and created several tables, one for each outcome category. They organized information to bring forward those mechanisms where patient portals contribute to outcomes and the variation in outcomes across different contexts.

9.3.6 Critical Reviews

Lastly, *critical reviews* aim to provide a critical evaluation and interpretive analysis of existing literature on a particular topic of interest to reveal strengths, weaknesses, contradictions, controversies, inconsistencies, and/or other important issues with respect to theories, hypotheses, research methods or results (Baumeister & Leary, 1997; Kirkevold, 1997). Unlike other review types, critical reviews attempt to take a reflective account of the research that has been done in a particular area of interest, and assess its credibility by using appraisal instruments or critical interpretive methods. In this way, critical reviews attempt to constructively inform other scholars about the weaknesses of prior research and strengthen knowledge development by giving focus and direction to studies for further improvement (Kirkevold, 1997).

Kitsiou, Paré, and Jaana (2013) provide an example of a critical review that assessed the methodological quality of prior systematic reviews of home telemonitoring studies for chronic patients. The authors conducted a comprehensive search on multiple databases to identify eligible reviews and subsequently used a validated instrument to conduct an in-depth quality appraisal. Results indicate that the majority of systematic reviews in this particular area suffer from important methodological flaws and biases that impair their internal validity and limit their usefulness for clinical and decision-making purposes. To

this end, they provide a number of recommendations to strengthen knowledge development towards improving the design and execution of future reviews on home telemonitoring.

9.4 Summary

Table 9.1 outlines the main types of literature reviews that were described in the previous sub-sections and summarizes the main characteristics that distinguish one review type from another. It also includes key references to methodological guidelines and useful sources that can be used by eHealth scholars and researchers for planning and developing reviews.

Table 9.1
Typology of Literature Reviews (adapted from Paré et al., 2015)

Review type	Overarching goal	Search strategy	Appraisal of included studies	Analysis and synthesis	Key references
Narrative review	Aims to summarize or synthesize what has been written on a particular topic but does not seek generalization or cumulative knowledge from what is reviewed.	Selective in nature. Authors usually select studies that support their own view.	No formal quality or risk of bias assessment of included primary studies is required.	Narrative using thematic analysis, chronological order, conceptual frameworks, content analysis or other classification criteria.	(Cronin et al., 2008; Green et al., 2006; Levy & Ellis, 2006; Webster & Watson, 2002)
Descriptive or mapping review	Seeks to identify interpretable patterns and gaps in the literature with respect to pre-existing propositions, theories, methodologies or findings.	Aims to identify a representative number of works on a particular topic. May or may not include comprehensive searching.	No formal quality or risk of bias assessment of included primary studies is required.	Quantitative or qualitative using descriptive statistics (e.g., frequencies), and content analysis methods.	(King & He, 2005; Paré et al., 2015; Petersen et al., 2015)

Table 9.1
Typology of Literature Reviews (adapted from Paré et al., 2015)

Review type	Overarching goal	Search strategy	Appraisal of included studies	Analysis and synthesis	Key references
Scoping review	Aims to provide an initial indication of potential size and scope of the extant research literature. May be conducted to identify nature and extent of research evidence, including ongoing research, with a view to determine the value of undertaking a full systematic review.	Comprehensive search using an iterative process that is guided by a requirement to identify all relevant literature (published and unpublished) suitable for answering the central research question regardless of study design. Uses explicit inclusion and exclusion criteria.	No formal quality or risk of bias assessment of included primary studies is required.	Uses analytic frameworks or thematic construction in order to present a narrative account of existing literature, as well as numerical analysis of the extent, nature and distribution of the studies included in the review.	(Arksey & O'Malley, 2005; Daudt et al., 2013; Levac et al., 2010).
Systematic review	Aims to aggregate, critically appraise, and synthesize in a single source all empirical evidence that meet a set of pre-specified eligibility criteria in order to answer in depth a clearly formulated research question to support evidence-based decision-making.	Exhaustive literature search of multiple sources and databases using highly sensitive and structured strategies to identify all available studies (published and unpublished) within resource limits that are eligible for inclusion. Uses a priori inclusion and exclusion criteria.	Two different quality assessments must be addressed in systematic reviews: (a) risk of bias in included studies, and (b) quality of evidence by outcome of interest. Both assessments require the use of validated instruments (e.g., Cochrane criteria and GRADE system).	Two different types of analyses and syntheses methods can be used: 1. Meta-analysis (statistical pooling of study results), and 2. qualitative/ narrative: use of vote counting, content analysis, frameworks, classification schemes, and/or tabulations.	(Borenstein et al., 2009; Higgins & Green, 2008; Liberati et al., 2009)

Table 9.1
Typology of Literature Reviews (adapted from Paré et al., 2015)

Review type	Overarching goal	Search strategy	Appraisal of included studies	Analysis and synthesis	Key references
Umbrella review	Tertiary type of evidence synthesis. Aims to compare and contrast findings from multiple systematic reviews in priority areas, at a variety of different levels, including different types of interventions for the same condition or alternatively, same interventions for different conditions, outcomes, problems, or populations and adverse effects.	Exhaustive literature search to identify all available systematic reviews (published and unpublished) within resource limits that are eligible for inclusion. No search for primary studies. Uses a priori inclusion and exclusion criteria.	Two different quality assessments must be addressed: (a) methodological quality assessment of the included systematic reviews, and (b) quality of evidence in included reviews. Both assessments require use of validated instruments (e.g., AMSTAR and GRADE system).	Many umbrella reviews will simply extract data from the underlying systematic reviews and summarize them in tables or figures. However, in some cases they may include indirect comparisons based on formal statistical analyses, especially if there is no evidence on direct comparisons.	(Becker & Oxman, 2008; Shea et al., 2009; Smith et al., 2011)
Realist review	Theory-driven interpretative review. Aims to inform, enhance, extend, or supplement conventional systematic reviews by including evidence from both quantitative and qualitative studies of complex interventions applied in diverse contexts to inform policy decision-making.	Can be systematic and comprehensive based on "a priori" criteria or iterative and purposive, aiming to provide a holistic interpretation of a phenomenon through theoretical saturation.	Quality or risk of bias assessment must be addressed using different instruments and/or frameworks for quantitative and qualitative studies. Questions about "quality" and "bias" are very different in the context of qualitative research.	Qualitative evidence synthesis. Can be aggregative or interpretive. Requires transparency. Can use content analysis, conceptual frameworks, as well as interpretive and mixed methods approaches.	(Pawson, 2006; Pawson et al., 2005; Whitlock et al., 2008)

Table 9.1
Typology of Literature Reviews (adapted from Paré et al., 2015)

Review type	Overarching goal	Search strategy	Appraisal of included studies	Analysis and synthesis	Key references
Critical review	Aims to provide a critical evaluation and interpretive analysis of existing literature on a particular topic of interest to reveal strengths, weaknesses, contradictions, controversies, inconsistencies, and/or other important issues with respect to theories, hypotheses, research methods or results.	Seeks to identify a representative number of articles that make the sample illustrative of the larger group of works in the field of study. May or may not include comprehensive searching.	No formal quality or risk of bias assessment of included primary studies is required.	Can apply a variety of analysis methods that can be grouped as either positivist (e.g., content analysis and frequencies) or interpretivist (e.g., meta-ethnography, critical interpretive synthesis) according to the authors' epistemological positions.	(Kirkevold, 1997; Paré et al., 2015)

Note. From "Synthesizing information systems knowledge: A typology of literature reviews," by G. Paré, M. C. Trudel, M. Jaana, and S. Kitsiou, 2015, *Information & Management, 52*(2), p. 187. Adapted with permission.

As shown in Table 9.1, each review type addresses different kinds of research questions or objectives, which subsequently define and dictate the methods and approaches that need to be used to achieve the overarching goal(s) of the review. For example, in the case of narrative reviews, there is greater flexibility in searching and synthesizing articles (Green et al., 2006). Researchers are often relatively free to use a diversity of approaches to search, identify, and select relevant scientific articles, describe their operational characteristics, present how the individual studies fit together, and formulate conclusions. On the other hand, systematic reviews are characterized by their high level of systematicity, rigour, and use of explicit methods, based on an "a priori" review plan that aims to minimize bias in the analysis and synthesis process (Higgins & Green, 2008). Some reviews are exploratory in nature (e.g., scoping/mapping reviews), whereas others may be conducted to discover patterns (e.g., descriptive reviews) or involve a synthesis approach that may include the critical analysis of prior research (Paré et al., 2015). Hence, in order to select the most appropriate type of review, it is critical to know before embarking on a review project, why the research synthesis is conducted and what type of methods are best aligned with the pursued goals.

9.5 Concluding Remarks

In light of the increased use of evidence-based practice and research generating stronger evidence (Grady et al., 2011; Lyden et al., 2013), review articles have become essential tools for summarizing, synthesizing, integrating or critically appraising prior knowledge in the eHealth field. As mentioned earlier, when rigorously conducted review articles represent powerful information sources for eHealth scholars and practitioners looking for state-of-the-art evidence. The typology of literature reviews we used herein will allow eHealth researchers, graduate students and practitioners to gain a better understanding of the similarities and differences between review types.

We must stress that this classification scheme does not privilege any specific type of review as being of higher quality than another (Paré et al., 2015). As explained above, each type of review has its own strengths and limitations. Having said that, we realize that the methodological rigour of any review — be it qualitative, quantitative or mixed — is a critical aspect that should be considered seriously by prospective authors. In the present context, the notion of rigour refers to the reliability and validity of the review process described in section 9.2. For one thing, *reliability* is related to the reproducibility of the review process and steps, which is facilitated by a comprehensive documentation of the literature search process, extraction, coding and analysis performed in the review. Whether the search is comprehensive or not, whether it involves a methodical approach for data extraction and synthesis or not, it is important that the review documents in an explicit and transparent manner the steps and approach that were used in the process of its development. Next, *validity* characterizes the degree to which the review process was conducted appropriately. It goes beyond documentation and reflects decisions related to the selection of the sources, the search terms used, the period of time covered, the articles selected in the search, and the application of backward and forward searches (vom Brocke et al., 2009). In short, the rigour of any review article is reflected by the explicitness of its methods (i.e., transparency) and the soundness of the approach used. We refer those interested in the concepts of rigour and quality to the work of Templier and Paré (2015) which offers a detailed set of methodological guidelines for conducting and evaluating various types of review articles.

To conclude, our main objective in this chapter was to demystify the various types of literature reviews that are central to the continuous development of the eHealth field. It is our hope that our descriptive account will serve as a valuable source for those conducting, evaluating or using reviews in this important and growing domain.

References

Ammenwerth, E., & de Keizer, N. (2004). An inventory of evaluation studies of information technology in health care. Trends in evaluation research, 1982-2002. *International Journal of Medical Informatics, 44*(1), 44–56.

Anderson, S., Allen, P., Peckham, S., & Goodwin, N. (2008). Asking the right questions: scoping studies in the commissioning of research on the organisation and delivery of health services. *Health Research Policy and Systems, 6*(7), 1–12. doi: 10.1186/1478-4505-6-7

Archer, N., Fevrier-Thomas, U., Lokker, C., McKibbon, K. A., & Straus, S. E. (2011). Personal health records: a scoping review. *Journal of American Medical Informatics Association, 18*(4), 515–522.

Arksey, H., & O'Malley, L. (2005). Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology, 8*(1), 19–32.

Bandara, W., Miskon, S., & Fielt, E. (2011). *A systematic, tool-supported method for conducting literature reviews in information systems.* Paper presented at the Proceedings of the 19th European Conference on Information Systems (ECIS 2011), June 9 to 11, Helsinki, Finland.

Baumeister, R. F., & Leary, M. R. (1997). Writing narrative literature reviews. *Review of General Psychology, 1*(3), 311–320.

Becker, L. A., & Oxman, A. D. (2008). Overviews of reviews. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions* (pp. 607–631). Hoboken, NJ: John Wiley & Sons, Ltd.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to meta-analysis*. Hoboken, NJ: John Wiley & Sons Inc.

Cook, D. J., Mulrow, C. D., & Haynes, B. (1997). Systematic reviews: Synthesis of best evidence for clinical decisions. *Annals of Internal Medicine, 126*(5), 376–380.

Cooper, H., & Hedges, L. V. (2009). Research synthesis as a scientific process. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 3–17). New York: Russell Sage Foundation.

Cooper, H. M. (1988). Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in Society, 1*(1), 104–126.

Cronin, P., Ryan, F., & Coughlan, M. (2008). Undertaking a literature review: a step-by-step approach. *British Journal of Nursing, 17*(1), 38–43.

Darlow, S., & Wen, K. Y. (2015). Development testing of mobile health interventions for cancer patient self-management: A review. *Health Informatics Journal* (online before print). doi: 10.1177/1460458215577994

Daudt, H. M., van Mossel, C., & Scott, S. J. (2013). Enhancing the scoping study methodology: a large, inter-professional team's experience with Arksey and O'Malley's framework. *BMC Medical Research Methodology, 13*, 48. doi: 10.1186/1471-2288-13-48

Davies, P. (2000). The relevance of systematic reviews to educational policy and practice. *Oxford Review of Education, 26*(3-4), 365–378.

Deeks, J. J., Higgins, J. P. T., & Altman, D. G. (2008). Analysing data and undertaking meta-analyses. In J. P. T. Higgins & S. Green (Eds.), *Cochrane handbook for systematic reviews of interventions* (pp. 243–296). Hoboken, NJ: John Wiley & Sons, Ltd.

Deshazo, J. P., Lavallie, D. L., & Wolf, F. M. (2009). Publication trends in the medical informatics literature: 20 years of "Medical Informatics" in MESH. *BMC Medical Informatics and Decision Making, 9*, 7. doi: 10.1186/1472-6947-9-7

Dixon-Woods, M., Agarwal, S., Jones, D., Young, B., & Sutton, A. (2005). Synthesising qualitative and quantitative evidence: a review of possible methods. *Journal of Health Services Research and Policy, 10*(1), 45–53.

Finfgeld-Connett, D., & Johnson, E. D. (2013). Literature search strategies for conducting knowledge-building and theory-generating qualitative systematic reviews. *Journal of Advanced Nursing, 69*(1), 194–204.

Grady, B., Myers, K. M., Nelson, E. L., Belz, N., Bennett, L., Carnahan, L., . . . Guidelines Working Group. (2011). Evidence-based practice for telemental health. *Telemedicine Journal and E Health, 17*(2), 131–148.

Green, B. N., Johnson, C. D., & Adams, A. (2006). Writing narrative literature reviews for peer-reviewed journals: secrets of the trade. *Journal of Chiropractic Medicine, 5*(3), 101–117.

Greenhalgh, T., Wong, G., Westhorp, G., & Pawson, R. (2011). Protocol–realist and meta-narrative evidence synthesis: evolving standards (RAMESES). *BMC Medical Research Methodology, 11*, 115.

Gurol-Urganci, I., de Jongh, T., Vodopivec-Jamsek, V., Atun, R., & Car, J. (2013). Mobile phone messaging reminders for attendance at healthcare appointments. *Cochrane Database System Review, 12*, CD007458. doi: 10.1002/14651858

Hart, C. (1998). *Doing a literature review: Releasing the social science research imagination*. London: SAGE Publications.

Higgins, J. P. T., & Green, S. (Eds.). (2008). *Cochrane handbook for systematic reviews of interventions: Cochrane book series*. Hoboken, NJ: Wiley-Blackwell.

Jesson, J., Matheson, L., & Lacey, F. M. (2011). *Doing your literature review: traditional and systematic techniques*. Los Angeles & London: SAGE Publications.

King, W. R., & He, J. (2005). Understanding the role and methods of meta-analysis in IS research. *Communications of the Association for Information Systems, 16*, 1.

Kirkevold, M. (1997). Integrative nursing research — an important strategy to further the development of nursing science and nursing practice. *Journal of Advanced Nursing, 25*(5), 977–984.

Kitchenham, B., & Charters, S. (2007). Guidelines for performing systematic literature reviews in software engineering. *EBSE Technical Report Version 2.3*. Keele & Durham, UK: Keele University & University of Durham.

Kitsiou, S., Paré, G., & Jaana, M. (2013). Systematic reviews and meta-analyses of home telemonitoring interventions for patients with chronic diseases: a critical assessment of their methodological quality. *Journal of Medical Internet Research, 15*(7), e150.

Kitsiou, S., Paré, G., & Jaana, M. (2015). Effects of home telemonitoring interventions on patients with chronic heart failure: an overview of systematic reviews. *Journal of Medical Internet Research, 17*(3), e63.

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: advancing the methodology. *Implementation Science, 5*(1), 69.

Levy, Y., & Ellis, T. J. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science, 9*, 181–211.

Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P. A., … Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *Annals of Internal Medicine, 151*(4), W-65.

Lyden, J. R., Zickmund, S. L., Bhargava, T. D., Bryce, C. L., Conroy, M. B., Fischer, G. S., . . . McTigue, K. M. (2013). Implementing health information technology in a patient-centered manner: Patient experiences with an online evidence-based lifestyle intervention. *Journal for Healthcare Quality, 35*(5), 47–57.

Mickan, S., Atherton, H., Roberts, N. W., Heneghan, C., & Tilson, J. K. (2014). Use of handheld computers in clinical practice: a systematic review. *BMC Medical Informatics and Decision Making, 14*, 56.

Moher, D. (2013). The problem of duplicate systematic reviews. *British Medical Journal, 347*(5040). doi: 10.1136/bmj.f5040

Montori, V. M., Wilczynski, N. L., Morgan, D., Haynes, R. B., & Hedges, T. (2003). Systematic reviews: a cross-sectional study of location and citation counts. *BMC Medicine, 1*, 2.

Mulrow, C. D. (1987). The medical review article: state of the science. *Annals of Internal Medicine, 106*(3), 485–488. doi:10.7326/0003-4819-106-3-485

Oates, B. J. (2011). Evidence-based information systems: A decade later. *Proceedings of the European Conference on Information Systems*. Retrieved from http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1221&context=ecis2011

Okoli, C., & Schabram, K. (2010). A guide to conducting a systematic literature review of information systems research. *SSRN Electronic Journal*.

Otte-Trojel, T., de Bont, A., Rundall, T. G., & van de Klundert, J. (2014). How outcomes are achieved through patient portals: a realist review. *Journal of American Medical Informatics Association*, *21*(4), 751–757.

Paré, G., Trudel, M.-C., Jaana, M., & Kitsiou, S. (2015). Synthesizing information systems knowledge: A typology of literature reviews. *Information & Management, 52*(2), 183–199.

Patsopoulos, N. A., Analatos, A. A., & Ioannidis, J. P. A. (2005). Relative citation impact of various study designs in the health sciences. *Journal of the American Medical Association, 293*(19), 2362–2366.

Paul, M. M., Greene, C. M., Newton-Dame, R., Thorpe, L. E., Perlman, S. E., McVeigh, K. H., & Gourevitch, M. N. (2015). The state of population health surveillance using electronic health records: A narrative review. *Population Health Management, 18*(3), 209–216.

Pawson, R. (2006). *Evidence-based policy: a realist perspective*. London: SAGE Publications.

Pawson, R., Greenhalgh, T., Harvey, G., & Walshe, K. (2005). Realist review—a new method of systematic review designed for complex policy interventions. *Journal of Health Services Research & Policy, 10*(Suppl 1), 21–34.

Petersen, K., Vakkalanka, S., & Kuzniarz, L. (2015). Guidelines for conducting systematic mapping studies in software engineering: An update. *Information and Software Technology, 64*, 1–18.

Petticrew, M., & Roberts, H. (2006). *Systematic reviews in the social sciences: A practical guide*: Malden, MA: Blackwell Publishing Co.

Rousseau, D. M., Manning, J., & Denyer, D. (2008). Evidence in management and organizational science: Assembling the field's full weight of scientific knowledge through syntheses. *The Academy of Management Annals, 2*(1), 475–515.

Rowe, F. (2014). What literature review is not: diversity, boundaries and recommendations. *European Journal of Information Systems, 23*(3), 241–255.

Shea, B. J., Hamel, C., Wells, G. A., Bouter, L. M., Kristjansson, E., Grimshaw, J., … Boers, M. (2009). AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. *Journal of Clinical Epidemiology, 62*(10), 1013–1020.

Shepperd, S., Lewin, S., Straus, S., Clarke, M., Eccles, M. P., Fitzpatrick, R., . . . Sheikh, A. (2009). Can we systematically review studies that evaluate complex interventions? *PLoS Medicine, 6*(8), e1000086.

Silva, B. M., Rodrigues, J. J., de la Torre Díez, I., López-Coronado, M., & Saleem, K. (2015). Mobile-health: A review of current state in 2015. *Journal of Biomedical Informatics, 56*, 265–272.

Smith, V., Devane, D., Begley, C., & Clarke, M. (2011). Methodology in conducting a systematic review of systematic reviews of healthcare interventions. *BMC Medical Research Methodology, 11*(1), 15.

Sylvester, A., Tate, M., & Johnstone, D. (2013). Beyond synthesis: re-presenting heterogeneous research literature. *Behaviour & Information Technology, 32*(12), 1199–1215.

Templier, M., & Paré, G. (2015). A framework for guiding and evaluating literature reviews. *Communications of the Association for Information Systems, 37*(6), 112–137.

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology, 8*(1), 45.

vom Brocke, J., Simons, A., Niehaves, B., Riemer, K., Plattfaut, R., & Cleven, A. (2009). *Reconstructing the giant: on the importance of rigour in documenting the literature search process.* Paper presented at the Proceedings of the 17th European Conference on Information Systems (ECIS 2009), Verona, Italy.

Webster, J., & Watson, R. T. (2002). Analyzing the past to prepare for the future: Writing a literature review. *Management Information Systems Quarterly, 26*(2), 11.

Whitlock, E. P., Lin, J. S., Chou, R., Shekelle, P., & Robinson, K. A. (2008). Using existing systematic reviews in complex systematic reviews. *Annals of Internal Medicine, 148*(10), 776–782.

Chapter 10

Methods for Comparative Studies

Francis Lau, Anne Holbrook

10.1 Introduction

In eHealth evaluation, comparative studies aim to find out whether group differences in eHealth system adoption make a difference in important outcomes. These groups may differ in their composition, the type of system in use, and the setting where they work over a given time duration. The comparisons are to determine whether significant differences exist for some predefined measures between these groups, while controlling for as many of the conditions as possible such as the composition, system, setting and duration.

According to the typology by Friedman and Wyatt (2006), comparative studies take on an objective view where events such as the use and effect of an eHealth system can be defined, measured and compared through a set of variables to prove or disprove a hypothesis. For comparative studies, the design options are experimental versus observational and prospective versus retrospective. The quality of eHealth comparative studies depends on such aspects of methodological design as the choice of variables, sample size, sources of bias, confounders, and adherence to quality and reporting guidelines.

In this chapter we focus on experimental studies as one type of comparative study and their methodological considerations that have been reported in the eHealth literature. Also included are three case examples to show how these studies are done.

10.2 Types of Comparative Studies

Experimental studies are one type of comparative study where a sample of participants is identified and assigned to different conditions for a given time duration, then compared for differences. An example is a hospital with two care

units where one is assigned a CPOE system to process medication orders electronically while the other continues its usual practice without a CPOE. The participants in the unit assigned to the CPOE are called the intervention group and those assigned to usual practice are the control group. The comparison can be performance or outcome focused, such as the ratio of correct orders processed or the occurrence of adverse drug events in the two groups during the given time period. Experimental studies can take on a randomized or non-randomized design. These are described below.

10.2.1 Randomized Experiments

In a randomized design, the participants are randomly assigned to two or more groups using a known randomization technique such as a random number table. The design is prospective in nature since the groups are assigned concurrently, after which the intervention is applied then measured and compared. Three types of experimental designs seen in eHealth evaluation are described below (Friedman & Wyatt, 2006; Zwarenstein & Treweek, 2009).

Randomized controlled trials (RCTs) – In RCTs participants are randomly assigned to an intervention or a control group. The randomization can occur at the patient, provider or organization level, which is known as the unit of allocation. For instance, at the patient level one can randomly assign half of the patients to receive EMR reminders while the other half do not. At the provider level, one can assign half of the providers to receive the reminders while the other half continues with their usual practice. At the organization level, such as a multisite hospital, one can randomly assign EMR reminders to some of the sites but not others.

Cluster randomized controlled trials (CRCTs) – In CRCTs, clusters of participants are randomized rather than by individual participant since they are found in naturally occurring groups such as living in the same communities. For instance, clinics in one city may be randomized as a cluster to receive EMR reminders while clinics in another city continue their usual practice.

Pragmatic trials – Unlike RCTs that seek to find out if an intervention such as a CPOE system works under ideal conditions, pragmatic trials are designed to find out if the intervention works under usual conditions. The goal is to make the design and findings relevant to and practical for decision-makers to apply in usual settings. As such, pragmatic trials have few criteria for selecting study participants, flexibility in implementing the intervention, usual practice as the comparator, the same compliance and follow-up

intensity as usual practice, and outcomes that are relevant to decision-makers.

10.2.2 Non-randomized Experiments

Non-randomized design is used when it is neither feasible nor ethical to randomize participants into groups for comparison. It is sometimes referred to as a quasi-experimental design. The design can involve the use of prospective or retrospective data from the same or different participants as the control group. Three types of non-randomized designs are described below (Harris et al., 2006).

Intervention group only with pretest and post-test design – This design involves only one group where a pretest or baseline measure is taken as the control period, the intervention is implemented, and a post-test measure is taken as the intervention period for comparison. For example, one can compare the rates of medication errors before and after the implementation of a CPOE system in a hospital. To increase study quality, one can add a second pretest period to decrease the probability that the pretest and post-test difference is due to chance, such as an unusually low medication error rate in the first pretest period. Other ways to increase study quality include adding an unrelated outcome such as patient case-mix that should not be affected, removing the intervention to see if the difference remains, and removing then re-implementing the intervention to see if the differences vary accordingly.

Intervention and control groups with post-test design – This design involves two groups where the intervention is implemented in one group and compared with a second group without the intervention, based on a post-test measure from both groups. For example, one can implement a CPOE system in one care unit as the intervention group with a second unit as the control group and compare the post-test medication error rates in both units over six months. To increase study quality, one can add one or more pretest periods to both groups, or implement the intervention to the control group at a later time to measure for similar but delayed effects.

Interrupted time series (ITS) design – In ITS design, multiple measures are taken from one group in equal time intervals, interrupted by the implementation of the intervention. The multiple pretest and post-test measures decrease the probability that the differences detected are due to chance or unrelated effects. An example is to take six consecutive monthly medication error rates as the pretest measures, implement the CPOE system, then take another

six consecutive monthly medication error rates as the post-test measures for comparison in error rate differences over 12 months. To increase study quality, one may add a concurrent control group for comparison to be more convinced that the intervention produced the change.

10.3 Methodological Considerations

The quality of comparative studies is dependent on their internal and external validity. Internal validity refers to the extent to which conclusions can be drawn correctly from the study setting, participants, intervention, measures, analysis and interpretations. External validity refers to the extent to which the conclusions can be generalized to other settings. The major factors that influence validity are described below.

10.3.1 Choice of Variables

Variables are specific measurable features that can influence validity. In comparative studies, the choice of dependent and independent variables and whether they are categorical and/or continuous in values can affect the type of questions, study design and analysis to be considered. These are described below (Friedman & Wyatt, 2006).

Dependent variables – This refers to outcomes of interest; they are also known as outcome variables. An example is the rate of medication errors as an outcome in determining whether CPOE can improve patient safety.

Independent variables – This refers to variables that can explain the measured values of the dependent variables. For instance, the characteristics of the setting, participants and intervention can influence the effects of CPOE.

Categorical variables – This refers to variables with measured values in discrete categories or levels. Examples are the type of providers (e.g., nurses, physicians and pharmacists), the presence or absence of a disease, and pain scale (e.g., 0 to 10 in increments of 1). Categorical variables are analyzed using non-parametric methods such as chi-square and odds ratio.

Continuous variables – This refers to variables that can take on infinite values within an interval limited only by the desired precision. Examples are blood pressure, heart rate and body temperature. Continuous variables are analyzed using parametric methods such as *t*-test, analysis of variance or multiple regression.

10.3.2 Sample Size

Sample size is the number of participants to include in a study. It can refer to patients, providers or organizations depending on how the unit of allocation is defined. There are four parts to calculating sample size. They are described below (Noordzij et al., 2010).

Significance level – This refers to the probability that a positive finding is due to chance alone. It is usually set at 0.05, which means having a less than 5% chance of drawing a false positive conclusion.

Power – This refers to the ability to detect the true effect based on a sample from the population. It is usually set at 0.8, which means having at least an 80% chance of drawing a correct conclusion.

Effect size – This refers to the minimal clinically relevant difference that can be detected between comparison groups. For continuous variables, the effect is a numerical value such as a 10-kilogram weight difference between two groups. For categorical variables, it is a percentage such as a 10% difference in medication error rates.

Variability – This refers to the population variance of the outcome of interest, which is often unknown and is estimated by way of standard deviation (SD) from pilot or previous studies for continuous outcome.

Table 10.1
Sample Size Equations for Comparing Two Groups with Continuous and Categorical Outcome Variables

Continuous variable Attributes	Categorical variable
n = 2[(a+b)2σ2]/(μ1-μ2)2 where	n = [(a+b)2(p1q1+p2q2)]/χ2
n = sample size for each group μ1 = population mean in group 1 μ2 = population mean in group 2 μ1- μ2 = desired difference between groups σ = population variance a = multiplier for significance level (or alpha) b = multiplier for power (or 1-beta)	n = sample size for each group p1 = proportion of participants with condition in group 1 q1 = proportion of participants without condition in group 1 p2 = proportion of participants with condition in group 2 q2 = proportion of participants without condition in group 2 χ = difference in outcome between two groups a = multiplier for significance level (or alpha) b = multiplier for power (or 1-beta)

An example of sample size calculation for an RCT to examine the effect of CDS on improving systolic blood pressure of hypertensive patients is provided

in the Appendix. Refer to the Biomath website from Columbia University (n.d.) for a simple Web-based sample size / power calculator.

10.3.3 Sources of Bias

There are five common sources of biases in comparative studies. They are selection, performance, detection, attrition and reporting biases (Higgins & Green, 2011). These biases, and the ways to minimize them, are described below (Vervloet et al., 2012).

Selection or allocation bias – This refers to differences between the composition of comparison groups in terms of the response to the intervention. An example is having sicker or older patients in the control group than those in the intervention group when evaluating the effect of EMR reminders. To reduce selection bias, one can apply randomization and concealment when assigning participants to groups and ensure their compositions are comparable at baseline.

Performance bias – This refers to differences between groups in the care they received, aside from the intervention being evaluated. An example is the different ways by which reminders are triggered and used within and across groups such as electronic, paper and phone reminders for patients and providers. To reduce performance bias, one may standardize the intervention and blind participants from knowing whether an intervention was received and which intervention was received.

Detection or measurement bias – This refers to differences between groups in how outcomes are determined. An example is where outcome assessors pay more attention to outcomes of patients known to be in the intervention group. To reduce detection bias, one may blind assessors from participants when measuring outcomes and ensure the same timing for assessment across groups.

Attrition bias – This refers to differences between groups in ways that participants are withdrawn from the study. An example is the low rate of participant response in the intervention group despite having received reminders for follow-up care. To reduce attrition bias, one needs to acknowledge the dropout rate and analyze data according to an intent-to-treat principle (i.e., include data from those who dropped out in the analysis).

Reporting bias – This refers to differences between reported and unreported findings. Examples include biases in publication, time lag, citation, language and outcome reporting depending on the nature and direction of the results. To reduce reporting bias, one may make the study protocol available with all pre-specified outcomes and report all expected outcomes in published results.

10.3.4 Confounders

Confounders are factors other than the intervention of interest that can distort the effect because they are associated with both the intervention and the outcome. For instance, in a study to demonstrate whether the adoption of a medication order entry system led to lower medication costs, there can be a number of potential confounders that can affect the outcome. These may include severity of illness of the patients, provider knowledge and experience with the system, and hospital policy on prescribing medications (Harris et al., 2006). Another example is the evaluation of the effect of an antibiotic reminder system on the rate of post-operative deep venous thromboses (DVTs). The confounders can be general improvements in clinical practice during the study such as prescribing patterns and post-operative care that are not related to the reminders (Friedman & Wyatt, 2006).

To control for confounding effects, one may consider the use of matching, stratification and modelling. Matching involves the selection of similar groups with respect to their composition and behaviours. Stratification involves the division of participants into subgroups by selected variables, such as comorbidity index to control for severity of illness. Modelling involves the use of statistical techniques such as multiple regression to adjust for the effects of specific variables such as age, sex and/or severity of illness (Higgins & Green, 2011).

10.3.5 Guidelines on Quality and Reporting

There are guidelines on the quality and reporting of comparative studies. The GRADE (Grading of Recommendations Assessment, Development and Evaluation) guidelines provide explicit criteria for rating the quality of studies in randomized trials and observational studies (Guyatt et al., 2011). The extended CONSORT (Consolidated Standards of Reporting Trials) Statements for non-pharmacologic trials (Boutron, Moher, Altman, Schulz, & Ravaud, 2008), pragmatic trials (Zwarestein et al., 2008), and eHealth interventions (Baker et al., 2010) provide reporting guidelines for randomized trials.

The GRADE guidelines offer a system of rating quality of evidence in systematic reviews and guidelines. In this approach, to support estimates of intervention effects RCTs start as high-quality evidence and observational studies as low-quality evidence. For each outcome in a study, five factors may rate down the quality of evidence. The final quality of evidence for each outcome would fall into one of high, moderate, low, and very low quality. These factors are listed below (for more details on the rating system, refer to Guyatt et al., 2011).

Design limitations – For RCTs they cover the lack of allocation concealment, lack of blinding, large loss to follow-up, trial stopped early or selective outcome reporting.

Inconsistency of results – Variations in outcomes due to unexplained heterogeneity. An example is the unexpected variation of effects across subgroups of patients by severity of illness in the use of preventive care reminders.

Indirectness of evidence – Reliance on indirect comparisons due to restrictions in study populations, intervention, comparator or outcomes. An example is the 30-day readmission rate as a surrogate outcome for quality of computer-supported emergency care in hospitals.

Imprecision of results – Studies with small sample size and few events typically would have wide confidence intervals and are considered of low quality.

Publication bias – The selective reporting of results at the individual study level is already covered under design limitations, but is included here for completeness as it is relevant when rating quality of evidence across studies in systematic reviews.

The original CONSORT Statement has 22 checklist items for reporting RCTs. For non-pharmacologic trials extensions have been made to 11 items. For pragmatic trials extensions have been made to eight items. These items are listed below. For further details, readers can refer to Boutron and colleagues (2008) and the CONSORT website (CONSORT, n.d.).

Title and abstract – one item on the means of randomization used.

Introduction – one item on background, rationale, and problem addressed by the intervention.

Methods – 10 items on participants, interventions, objectives, outcomes, sample size, randomization (sequence generation, allocation concealment, implementation), blinding (masking), and statistical methods.

Results – seven items on participant flow, recruitment, baseline data, numbers analyzed, outcomes and estimation, ancillary analyses, adverse events.

Discussion – three items on interpretation, generalizability, overall evidence.

The CONSORT Statement for eHealth interventions describes the relevance of the CONSORT recommendations to the design and reporting of eHealth studies with an emphasis on Internet-based interventions for direct use by patients, such as online health information resources, decision aides and PHRs. Of particular importance is the need to clearly define the intervention components, their role in the overall care process, target population, implementation process, primary and secondary outcomes, denominators for outcome analyses, and real world potential (for details refer to Baker et al., 2010).

10.4 Case Examples

10.4.1 Pragmatic RCT in Vascular Risk Decision Support

Holbrook and colleagues (2011) conducted a pragmatic RCT to examine the effects of a CDS intervention on vascular care and outcomes for older adults. The study is summarized below.

Setting – Community-based primary care practices with EMRs in one Canadian province.

Participants – English-speaking patients 55 years of age or older with diagnosed vascular disease, no cognitive impairment and not living in a nursing home, who had a provider visit in the past 12 months.

Intervention – A Web-based individualized vascular tracking and advice CDS system for eight top vascular risk factors and two diabetic risk factors, for use by both providers and patients and their families. Providers and staff could update the patient's profile at any time and the CDS algorithm ran nightly to update recommendations and colour highlighting used in the tracker interface. Intervention patients had Web access to the tracker, a print version mailed to them prior to the visit, and telephone support on advice.

Design – Pragmatic, one-year, two-arm, multicentre RCT, with randomization upon patient consent by phone, using an allocation-concealed online program. Randomization was by patient with stratification by provider using a block size of six. Trained reviewers examined EMR data and conducted patient telephone interviews to collect risk factors, vascular history, and vascular events. Providers completed questionnaires on the intervention at study

end. Patients had final 12-month lab checks on urine albumin, low-density lipoprotein cholesterol, and A1c levels.

Outcomes – Primary outcome was based on change in process composite score (PCS) computed as the sum of frequency-weighted process score for each of the eight main risk factors with a maximum score of 27. The process was considered met if a risk factor had been checked. PCS was measured at baseline and study end with the difference as the individual primary outcome scores. The main secondary outcome was a clinical composite score (CCS) based on the same eight risk factors compared in two ways: a comparison of the mean number of clinical variables on target and the percentage of patients with improvement between the two groups. Other secondary outcomes were actual vascular event rates, individual PCS and CCS components, ratings of usability, continuity of care, patient ability to manage vascular risk, and quality of life using the EuroQol five dimensions questionnaire (EQ-5D).

Analysis – 1,100 patients were needed to achieve 90% power in detecting a one-point PCS difference between groups with a standard deviation of five points, two-tailed *t*-test for mean difference at 5% significance level, and a withdrawal rate of 10%. The PCS, CCS and EQ-5D scores were analyzed using a generalized estimating equation accounting for clustering within providers. Descriptive statistics and $\chi 2$ tests or exact tests were done with other outcomes.

Findings – 1,102 patients and 49 providers enrolled in the study. The intervention group with 545 patients had significant PCS improvement with a difference of 4.70 ($p < .001$) on a 27-point scale. The intervention group also had significantly higher odds of rating improvements in their continuity of care (4.178, $p < .001$) and ability to improve their vascular health (3.07, $p < .001$). There was no significant change in vascular events, clinical variables and quality of life. Overall the CDS intervention led to reduced vascular risks but not to improved clinical outcomes in a one-year follow-up.

10.4.2 Non-randomized Experiment in Antibiotic Prescribing in Primary Care

Mainous, Lambourne, and Nietert (2013) conducted a prospective non-randomized trial to examine the impact of a CDS system on antibiotic prescribing for acute respiratory infections (ARIS) in primary care. The study is summarized below.

Setting – A primary care research network in the United States whose members use a common EMR and pool data quarterly for quality improvement and research studies.

Participants – An intervention group with nine practices across nine states, and a control group with 61 practices.

Intervention – Point-of-care CDS tool as customizable progress note templates based on existing EMR features. CDS recommendations reflect Centre for Disease Control and Prevention (CDC) guidelines based on a patient's predominant presenting symptoms and age. CDS was used to assist in ARI diagnosis, prompt antibiotic use, record diagnosis and treatment decisions, and access printable patient and provider education resources from the CDC.

Design – The intervention group received a multi-method intervention to facilitate provider CDS adoption that included quarterly audit and feedback, best practice dissemination meetings, academic detailing site visits, performance review and CDS training. The control group did not receive information on the intervention, the CDS or education. Baseline data collection was for three months with follow-up of 15 months after CDS implementation.

Outcomes – The outcomes were frequency of inappropriate prescribing during an ARI episode, broad-spectrum antibiotic use and diagnostic shift. Inappropriate prescribing was computed by dividing the number of ARI episodes with diagnoses in the inappropriate category that had an antibiotic prescription by the total number of ARI episodes with diagnosis for which antibiotics are inappropriate. Broad-spectrum antibiotic use was computed by all ARI episodes with a broad-spectrum antibiotic prescription by the total number of ARI episodes with an antibiotic prescription. Antibiotic drift was computed in two ways: dividing the number of ARI episodes with diagnoses where antibiotics are appropriate by the total number of ARI episodes with an antibiotic prescription; and dividing the number of ARI episodes where antibiotics were inappropriate by the total number of ARI episodes. Process measure included frequency of CDS template use and whether the outcome measures differed by CDS usage.

Analysis – Outcomes were measured quarterly for each practice, weighted by the number of ARI episodes during the quarter to assign greater weight to practices with greater numbers of relevant episodes and to periods with greater numbers of relevant episodes.

Weighted means and 95% CIs were computed separately for adult and pediatric (less than 18 years of age) patients for each time period for both groups. Baseline means in outcome measures were compared between the two groups using weighted independent-sample *t*-tests. Linear mixed models were used to compare changes over the 18-month period. The models included time, intervention status, and were adjusted for practice characteristics such as specialty, size, region and baseline ARIs. Random practice effects were included to account for clustering of repeated measures on practices over time. *P*-values of less than 0.05 were considered significant.

Findings – For adult patients, inappropriate prescribing in ARI episodes declined more among the intervention group (-0.6%) than the control group (4.2%)($p = 0.03$), and prescribing of broad-spectrum antibiotics declined by 16.6% in the intervention group versus an increase of 1.1% in the control group ($p < 0.0001$). For pediatric patients, there was a similar decline of 19.7% in the intervention group versus an increase of 0.9% in the control group ($p < 0.0001$). In summary, the CDS had a modest effect in reducing inappropriate prescribing for adults, but had a substantial effect in reducing the prescribing of broad-spectrum antibiotics in adult and pediatric patients.

10.4.3 Interrupted Time Series on EHR Impact in Nursing Care

Dowding, Turley, and Garrido (2012) conducted a prospective ITS study to examine the impact of EHR implementation on nursing care processes and outcomes. The study is summarized below.

Setting – Kaiser Permanente (KP) as a large not-for-profit integrated healthcare organization in the United States.

Participants – 29 KP hospitals in the northern and southern regions of California.

Intervention – An integrated EHR system implemented at all hospitals with CPOE, nursing documentation and risk assessment tools. The nursing component for risk assessment documentation of pressure ulcers and falls was consistent across hospitals and developed by clinical nurses and informaticists by consensus.

Design – ITS design with monthly data on pressure ulcers and quarterly data on fall rates and risk collected over seven years be-

tween 2003 and 2009. All data were collected at the unit level for each hospital.

Outcomes – Process measures were the proportion of patients with a fall risk assessment done and the proportion with a hospital-acquired pressure ulcer (HAPU) risk assessment done within 24 hours of admission. Outcome measures were fall and HAPU rates as part of the unit-level nursing care process and nursing sensitive outcome data collected routinely for all California hospitals. Fall rate was defined as the number of unplanned descents to the floor per 1,000 patient days, and HAPU rate was the percentage of patients with stages I-IV or unstageable ulcer on the day of data collection.

Analysis – Fall and HAPU risk data were synchronized using the month in which the EHR was implemented at each hospital as time zero and aggregated across hospitals for each time period. Multivariate regression analysis was used to examine the effect of time, region and EHR.

Findings – The EHR was associated with significant increase in document rates for HAPU risk (2.21; 95% CI 0.67 to 3.75) and non-significant increase for fall risk (0.36; -3.58 to 4.30). The EHR was associated with 13% decrease in HAPU rates (-0.76; -1.37 to -0.16) but no change in fall rates (-0.091; -0.29 to 011). Hospital region was a significant predictor of variation for HAPU (0.72; 0.30 to 1.14) and fall rates (0.57; 0.41 to 0.72). During the study period, HAPU rates decreased significantly (-0.16; -0.20 to -0.13) but not fall rates (0.0052; -0.01 to 0.02). In summary, EHR implementation was associated with a reduction in the number of HAPUs but not patient falls, and changes over time and hospital region also affected outcomes.

10.5 Summary

In this chapter we introduced randomized and non-randomized experimental designs as two types of comparative studies used in eHealth evaluation. Randomization is the highest quality design as it reduces bias, but it is not always feasible. The methodological issues addressed include choice of variables, sample size, sources of biases, confounders, and adherence to reporting guidelines. Three case examples were included to show how eHealth comparative studies are done.

References

Baker, T. B., Gustafson, D. H., Shaw, B., Hawkins, R., Pingree, S., Roberts, L., & Strecher, V. (2010). Relevance of CONSORT reporting criteria for research on eHealth interventions. *Patient Education and Counselling, 81*(suppl. 7), 77–86.

Columbia University. (n.d.). Statistics: sample size / power calculation. *Biomath* (Division of Biomathematics/Biostatistics), Department of Pediatrics. New York: Columbia University Medical Centre. Retrieved from http://www.biomath.info/power/index.htm

Boutron, I., Moher, D., Altman, D. G., Schulz, K. F., Ravaud, P., & CONSORT Group. (2008). Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: Explanation and elaboration. *Annals of Internal Medicine, 148*(4), 295–309.

Cochrane Collaboration. (n.d.). *Cochrane handbook*. London: Author. Retrieved from http://handbook.cochrane.org/

CONSORT Group. (n.d.). *The CONSORT statement*. Retrieved from http://www.consort-statement.org/

Dowding, D. W., Turley, M., & Garrido, T. (2012). The impact of an electronic health record on nurse sensitive patient outcomes: an interrupted time series analysis. *Journal of the American Medical Informatics Association, 19*(4), 615–620.

Friedman, C. P., & Wyatt, J. C. (2006). *Evaluation methods in biomedical informatics* (2nd ed.). New York: Springer Science + Business Media, Inc.

Guyatt, G., Oxman, A. D., Akl, E. A., Kunz, R., Vist, G., Brozek, J., … Schunemann, H. J. (2011). GRADE guidelines: 1. Introduction – GRADE evidence profiles and summary of findings tables. *Journal of Clinical Epidemiology, 64*(4), 383–394.

Harris, A. D., McGregor, J. C., Perencevich, E. N., Furuno, J. P., Zhu, J., Peterson, D. E., & Finkelstein, J. (2006). The use and interpretation of quasi-experimental studies in medical informatics. *Journal of the American Medical Informatics Association, 13*(1), 16–23.

Higgins, J. P. T., & Green, S. (Eds.). (2011). *Cochrane handbook for systematic reviews of interventions* (Version 5.1.0, updated March 2011). London: The Cochrane Collaboration. Retrieved from http://handbook.cochrane.org/

Holbrook, A., Pullenayegum, E., Thabane, L., Troyan, S., Foster, G., Keshavjee, K., … Curnew, G. (2011). Shared electronic vascular risk decision support in primary care. Computerization of medical practices for the enhancement of therapeutic effectiveness (COMPETE III) randomized trial. *Archives of Internal Medicine, 171*(19), 1736–1744.

Mainous III, A. G., Lambourne, C. A., & Nietert, P. J. (2013). Impact of a clinical decision support system on antibiotic prescribing for acute respiratory infections in primary care: quasi-experimental trial. *Journal of the American Medical Informatics Association*, *20*(2), 317–324.

Noordzij, M., Tripepi, G., Dekker, F. W., Zoccali, C., Tanck, M. W., & Jager, K. J. (2010). Sample size calculations: basic principles and common pitfalls. *Nephrology Dialysis Transplantation, 25*(5), 1388–1393. Retrieved from http://ndt.oxfordjournals.org/content/early/2010/01/12/ndt.gfp732.short

Vervloet, M., Linn, A. J., van Weert, J. C. M., de Bakker, D. H., Bouvy, M. L., & van Dijk, L. (2012). The effectiveness of interventions using electronic reminders to improve adherence to chronic medication: A systematic review of the literature. *Journal of the American Medical Informatics Association*, *19*(5), 696–704.

Zwarenstein, M., Treweek, S., Gagnier, J. J., Altman, D. G., Tunis, S., Haynes, B., Oxman, A. D., & Moher, D., for the CONSORT and Pragmatic Trials in Healthcare (Practihc) groups. (2008). Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *British Medical Journal, 337*, a2390. doi: 10.1136/bmj.a2390

Zwarenstein, M., & Treweek, S. (2009). What kind of randomized trials do we need? *Canadian Medical Association Journal, 180*(10), 998–1000.

Appendix

Example of Sample Size Calculation

This is an example of sample size calculation for an RCT that examines the effect of a CDS system on reducing systolic blood pressure in hypertensive patients. The case is adapted from the example described in the publication by Noordzij et al. (2010).

(a) Systolic blood pressure as a continuous outcome measured in mmHg

Based on similar studies in the literature with similar patients, the systolic blood pressure values from the comparison groups are expected to be normally distributed with a standard deviation of 20 mmHg. The evaluator wishes to detect a clinically relevant difference of 15 mmHg in systolic blood pressure as an outcome between the intervention group with CDS and the control group without CDS. Assuming a significance level or alpha of 0.05 for 2-tailed *t*-test and power of 0.80, the corresponding multipliers[1] are 1.96 and 0.842, respectively. Using the sample size equation for continuous outcome below we can calculate the sample size needed for the above study.

n = 2[(a+b)2σ2]/(μ1-μ2)2 where
n = sample size for each group
μ1 = population mean of systolic blood pressures in intervention group
μ2 = population mean of systolic blood pressures in control group
μ1- μ2 = desired difference in mean systolic blood pressures between groups
σ = population variance
a = multiplier for significance level (or alpha)
b = multiplier for power (or 1-beta)

Providing the values in the equation would give the sample size (n) of 28 samples per group as the result
n = 2[(1.96+0.842)2(202)]/152 or 28 samples per group

(b) Systolic blood pressure as a categorical outcome measured as below or above 140 mmHg (i.e., hypertension yes/no)

In this example a systolic blood pressure from a sample that is above 140 mmHg is considered an event of the patient with hypertension. Based on published literature the proportion of patients in the general population with hypertension is 30%. The evaluator wishes to detect a clinically relevant difference

1 From Table 3 on p. 1392 of Noordzij et al. (2010).

of 10% in systolic blood pressure as an outcome between the intervention group with CDS and the control group without CDS. This means the expected proportion of patients with hypertension is 20% (p1 = 0.2) in the intervention group and 30% (p2 = 0.3) in the control group. Assuming a significance level or alpha of 0.05 for 2-tailed *t*-test and power of 0.80 the corresponding multipliers are 1.96 and 0.842, respectively. Using the sample size equation for categorical outcome below, we can calculate the sample size needed for the above study.

n = [(a+b)2(p1q1+p2q2)]/χ2
n = sample size for each group
p1 = proportion of patients with hypertension in intervention group
q1 = proportion of patients without hypertension in intervention group (or 1-p1)
p2 = proportion of patients with hypertension in control group
q2 = proportion of patients without hypertension in control group (or 1-p2)
χ = desired difference in proportion of hypertensive patients between two groups
a = multiplier for significance level (or alpha)
b = multiplier for power (or 1-beta)

Providing the values in the equation would give the sample size (n) of 291 samples per group as the result
n = [(1.96+0.842)2((0.2)(0.8)+(0.3)(0.7))]/(0.1)2 or 291 samples per group

Chapter 11

Methods for Descriptive Studies

Yulong Gu, Jim Warren

11.1 Introduction

Descriptive studies in eHealth evaluations aim to assess the success of eHealth systems in terms of the system planning, design, implementation, use and impact. Descriptive studies focus on describing the process and impact of eHealth system development and implementation, which often are contextualized within the implementation environment (e.g., a healthcare organization). The descriptive nature of the evaluation design distinguishes descriptive studies from comparative studies such as a before/after study or a randomized controlled trial. In a 2003 literature review on evaluations of inpatient clinical information systems by van der Meijden and colleagues, four types of study design were identified: correlational, comparative, descriptive, and case study (van der Meijden, Tange, Troost, & Hasman, 2003). This review inherited the distinction between objectivist and subjectivist studies described by Friedman and Wyatt (1997); in the review, van der Meijden and colleagues defined descriptive study as an objectivist study to measure outcome variable(s) against predefined requirements, and case study as an subjectivist study of a phenomenon in its natural context using data from multiple sources — quantitatively or qualitatively (van der Meijden et al., 2003). For simplicity, we include case study under the descriptive study category in this chapter, and promote methodological components of qualitative, quantitative, and mixed methods for designing eHealth evaluations in this category. Adopting this wider scope, the following sections introduce the types of descriptive studies in eHealth evaluations, address methodological considerations, and provide examples of such studies.

11.2 Types of Descriptive Studies

There are five main types of descriptive studies undertaken in eHealth evaluations. These are separated by the overall study design and the methods of data collection and analysis, as well as by the objectives and assumptions of the evaluation. The five types can be termed: qualitative studies, case studies, usability studies, mixed methods studies, and other methods studies (including ethnography, action research, and grounded theory studies).

11.2.1 Qualitative Studies

The methodological approach of qualitative studies for eHealth evaluations is particularly appropriate when "we are interested in the 'how' or 'why' of processes and people using technology" (McKibbon, 2015). Qualitative study design can be used in both formative and summative evaluations of eHealth interventions. The qualitative methods of data collection and analysis include observation, documentation, interview, focus group, and open-ended questionnaire. These methods help understand the experiences of people using or planning on using eHealth solutions.

In qualitative studies, an interpretivist view is often adopted. This means qualitative researchers start from the position that their knowledge of reality is a social construction by human actors; their theories concerning reality are ways of making sense of the world, and shared meanings are a form of intersubjectivity rather than objectivity (Walsham, 2006). There is also increasing uptake of critical theory and critical realism in qualitative health evaluation research (McEvoy & Richards, 2003). The assumption for this paradigm is that reality exists independent of the human mind regardless of whether it can be comprehended or directly experienced (Levers, 2013). Irrespective of the different epistemological assumptions, qualitative evaluations of eHealth interventions apply similar data collection tools and analysis techniques to describe, interpret, and challenge people's perceptions and experiences with the environment where the intervention has been implemented or is being planned for implementation.

11.2.2 Case Studies

A case study investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident (Yin, 2011). Case study methods are commonly used in social sciences, and increasingly in information systems (IS) research since the 1980s, to produce meaningful results from a holistic investigation into the complex and ubiquitous interactions among organizations, technologies, and people (Dubé & Paré, 2003). The key decisions in designing a case study involve: (a) how to define the case being studied; (b) how to determine the relevant data to be collected; and (c) what should be done with the data once collected (Yin, 2011). These decisions remain the crucial questions to ask when designing an eHealth evaluation case study. In eHealth evaluations, the fundamental question regarding the case definition is often answered based on consultation with a

range of eHealth project stakeholders. Investigations should also be undertaken at an early stage in the case study design into the availability of qualitative data sources — whether informants or documents — as well as the feasibility of collecting quantitative data. For instance, eHealth systems often leave digital footprints in the form of system usage patterns and user profiles which may help in assessing system uptake and potentially in understanding system impact.

Case study design is versatile and flexible; it can be used with any philosophical perspective (e.g., positivist, interpretivist, or critical); it can also combine qualitative and quantitative data collection methods (Dubé & Paré, 2003). Case study research can involve a single case study or multiple case studies; and can take the strategy of an explanatory, exploratory or descriptive approach (Yin, 2011). The quality of eHealth evaluation case studies relies on choosing appropriate study modes according to the purpose and context of the evaluation. This context should also be described in detail in the study reporting; this will assist with demonstrating the credibility and generalizability of the research results (Benbasat, Goldstein, & Mead, 1987; Yin, 2011).

11.2.3 Usability Studies

Usability of an information system refers to the capacity of the system to allow users to carry out their tasks safely, effectively, efficiently and enjoyably (Kushniruk & Patel, 2004; Preece, Rogers, & Sharp, 2002; Preece et al., 1994). Kushniruk and Patel (2004) categorized the usability studies that involve user representatives as usability testing studies and the expert-based studies as usability inspection studies. They highlighted heuristic evaluation (Nielsen & Molich, 1990) and cognitive walkthrough (Polson, Lewis, Rieman, & Wharton, 1992) as two useful expert-based usability inspection approaches. Usability studies can evaluate an eHealth system in terms of both the design and its implementation. The goals of usability evaluations include assessing the extent of system functionality, the effect of interface on users, and identifying specific problems. Usability testing should be considered in all stages of the system design life cycle. The idea of testing early and often is a valuable principle for having a good usable system (e.g., to get usability evaluation results from early-stage prototypes including paper prototypes). Another principle, although challenging for eHealth innovations, is to involve users early and often — that is, to keep real users close to the design process. The interaction design model (Cooper, 2004) recommends having at least one user as part of the design team from the beginning, so that right from the formulation of the product its concept actually makes sense to the type of users it's aimed for; and the users themselves should participate in the usability testing.

A classic usability study is done through user participation, either in a laboratory setting or in the natural environment. There is also a suite of techniques that are sometimes called "discount" usability testing or expert-based evaluation (as they are applied by usability experts rather than end users). The most prominent expert-based approach is heuristic evaluation (Nielsen & Molich, 1990).

Whichever approach is taken for usability studies, the target measures for usability are similar:

- How long is it taking users to do the task?
- How accurate are users in doing the task?
- How long does it take users to learn to do the task with the system?
- How well do users remember how to use the system from earlier sessions?
- And, in general, how happy are users about having worked the task with the tool?

A usability specification can combine these five measures into requirements, such as: at least 90% of users can perform a given task correctly within no more than five minutes one week after completing a 30-minute tutorial.

11.2.4 Mixed Methods Studies

Increasing uptake and recognition of mixed methods studies, which combines qualitative and quantitative components in one research study, have been observed in health sciences and health services research (Creswell, Klassen, Plano, & Smith, 2011; Wisdom, Cavaleri, Onwuegbuzie, & Green, 2012). Mixed methods studies draw on the strength of utilizing multiple methods, but have challenges inherent to the approach as well, such as how to justify diverse philosophical positions and multiple theoretical frameworks, and how to integrate multiple forms of data. A key element in reporting mixed methods studies is to describe the study procedures in detail to inform readers about the study quality.

Given the nature of eHealth innovations — often new, complex and hard to measure — a mixed methods design is particularly suitable for their evaluations to collect robust evidence on not only their effectiveness, but also the real-life contextual understandings of their implementation. For instance, the system transactional data may indicate the technology uptake and usage pattern; and end user interviews collect people's insights into why they think certain events have happened and how to do things better.

11.2.5 Other Methods (ethnography, action research, grounded theory)

In addition to the above four main categories of designs used in eHealth evaluation studies, this section introduces a few other relevant and powerful approaches, including ethnography, action research, and grounded theory methods.

- With origins in anthropology, an ethnographic approach to information systems research aims to provide rich insights into the human, social and organizational aspects of systems development and application (Harvey & Myers, 1995). A distinguishing feature of ethnographic research is participant observation, that is, the researcher must have been there and "lived" there for reasonable length of time (Myers, 1997a). Interviews, surveys, and field notes can also be used in ethnography studies to collect data.

- Similarly, multiple data collection methods can be used in an action research study. The key feature of action research design is its "participatory, democratic process concerned with developing practical knowing" (Reason & Bradbury, 2001, p. 1). Action research studies naturally mix the problem-solving activities with research activities to produce knowledge (Chiasson, Germonprez, & Mathiassen, 2009), and often take an iterative process of planning, acting, observing, and reflecting (McNiff & Whitehead, 2002).

- Grounded theory is defined as an inductive methodology to generate theories through a rigorous research process leading to the emergence of conceptual categories; and these concepts as categories are related to each other as a theoretical explanation of the actions that continually resolve the main concern of the participants in a substantive area (Glaser & Strauss, 1967; Rhine, 2008). In the field of information systems research, grounded theory methodology is useful for developing context-based, process-oriented descriptions and explanations of the phenomena (Myers, 1997b). A 2013 review found that the most common use of grounded theory in Information Systems studies is the application of grounded theory techniques, typically for data analysis purposes (Matavire & Brown, 2013).

It is worth noting that the use of the above methods does not exclude other designs. For instance, ethnographic observations can be undertaken as one element in a mixed methods case study (Greenhalgh, Hinder, Stramer, Bratan, & Russell, 2010).

11.3 Methodological Considerations

There are a range of methodological issues that need to be considered when designing, undertaking and reporting a descriptive eHealth evaluation. These issues may emerge throughout the study procedures, from defining study objectives to presenting data interpretation. This section provides a quick guide

for addressing the most critical issues in order to choose and describe an appropriate approach in your study.

11.3.1 Study Objectives and Questions

The high-level goals of an eHealth evaluation study are often planned in the initial phase of the study. The goals define what the study is meant to reveal and what is to be learned. These may be documented as a multilevel statement of high-level intentions or questions. This statement is then expanded in the methodology section of the final study report with specific aspects of the purpose of the evaluation: that is, things you want to find out. For instance, if the innovation were an electronic referral (e-referral) system:

- The acceptance of e-referrals by all impacted healthcare workers.
- The impact of e-referrals on safety, efficiency and timeliness of healthcare delivery.
- The key problems and issues emerging from a technical and management perspective in implementation of e-referrals.

Some of the above specific statements may be expressed as testable hypotheses; for example, "Use of e-referrals is widely accepted by General Practitioners (GPs)." A good use of expanded objectives is to state specific research questions; for example, we might ask, "Do GPs prefer e-referrals to hard copy referrals?" as part of the "acceptance" assessment objective above.

11.3.2 Observable and Contextual Variables

In many cases, eHealth evaluation will be linked to (as part of, or coming after) a health IS implementation project that had a business case based on specific expected benefits of the technology, and specific functional and non-functional requirements as critical success factors of the project. These should be part of the evaluation's benefits framework. International literature (e.g., the benefits found with similar technology when evaluated overseas) may also inform the framework. The establishment of benefits framework in an eHealth evaluation will dictate the study design and variables selection, as well as the methods of data collection and analysis. For instance, observable variables to measure system outcome may include: mortality, morbidity, readmission, length of stay, patient functional status or quality of health/life.

One of the strengths of descriptive studies is that the study findings are contextualized within the system implementation environment. Hence, it is a good practice to explain in the methodology what system(s) is evaluated, including the technologies introduced, years and geography of implementation and use, as well as the healthcare delivery organizations and user groups involved in their use. Contextual variables also include those detailing the evaluation parameters

such as research study period and those contextual conditions that are relevant to the system implementation success or failure, for example, organizational structure and funding model.

11.3.3 Credibility, Authenticity and Contextualization

The philosophy of evaluation that is taken along with the detailed research procedures should be described to demonstrate the study rigour, reliability, validity and credibility. The methods used should also be detailed (e.g., interviews of particular user or management groups, analysis of particular data files, statistical procedures, etc.). Data triangulation (examining the consistency of different data sources) is a common technique to enhance the research quality. Where any particularly novel methods are used, they should be explained with reference to academic literature and/or particular projects from which they have arisen; ideally, they should be justified with comparison to other methods that suit similar purposes.

Authenticity is regarded as a feature particular to naturalistic inquiry (and ethnographic naturalism), an approach to inquiry that aims to generate a genuine or true understanding of people's experiences (Schwandt, 2007). In a wider sense of descriptive eHealth evaluation studies, it is important to maintain research authenticity — to convey a genuine understanding of the project stakeholders' experiences from their own point of view.

Related to the above discussion on credibility and authenticity, the goal of contextualizing study findings is to support the final theory by seeing whether "the meaning system and rules of behaviour make sense to those being studied" (Neuman, 2003). For example, to draw a "rich picture" of the impact of the evaluated eHealth implementation, the study may inquire and report on "How has it impacted the social context (e.g., communications, perceived roles and responsibilities, and how the users feel about themselves and others)?"

11.3.4 Theoretical Sampling and Saturation

Theoretical sampling is an important tool in grounded theory studies. It is to decide, on analytic grounds, what data to collect next and where to find them (Glaser & Strauss, 1967). This requires calculation and imagination from the analyst in order to move the theory along quickly and efficiently. The basic criterion is to govern the selection of comparison groups for discovering theory based on their theoretical relevance for furthering the development of emerging categories (Glaser & Strauss, 1967).

In studies that collect data via interviews, ideally the interviewing should continue, extending with further theoretical sampling, until the evaluators have reached "saturation" — the point where all the relevant contributions from new interviewees neatly fit categories identified from earlier interviews. Often time and budget do not allow full saturation, in which cases the key topics of interest and major data themes need to be confirmed, for example, by repeating emphasis from individuals in similar roles.

11.3.5 Data Collection and Analysis

Descriptive studies may use a range of diverse and flexible methods in data collection and analysis. Detailed description of the data collection methods used will help readers understand exactly how the study achieves the measurements that are relevant to your approach and measurement criteria. This includes how interviewees are identified, and sources of documents and electronic data, as well as pre-planned interview questions and questionnaires.

In terms of describing quantitative data analysis methods, all statistical procedures associated with the production of quantitative results need to be stated. Similarly, all analysis protocols for qualitative data should be clarified (e.g., the data coding methods used).

11.3.6 Interpretation and Dissemination

Key findings from descriptive studies should provide answers to the research objectives/questions. In general, these findings can be tabulated against the benefits framework you introduced as part of the methodology. Interpretation of the findings may characterize how the eHealth intervention enabled a transformation in healthcare practices. Moreover, when explaining the interpretation and implications drawn from the evaluation results, the key implications can be organized into formal recommendations.

In terms of evaluation dissemination, the study findings should reach all stakeholders considering uptake of similar technology. Evaluation and dissemination as iterative cycles should be considered. Feedback from dissemination of interim findings is a valuable component of the evaluation per se. A dissemination strategy should be planned, specifying the dissemination time frame and pathways (e.g., conventional written reporting, face-to-face reporting, Web 2.0, commercial media and academic publications).

11.4 Exemplary Cases

This section illustrates two descriptive eHealth evaluation studies, one case study as part of the commissioned evaluation on the implementation and impact of the summary care record (SCR) and HealthSpace programmes in the United Kingdom, and the other study from Canada as a usability evaluation to inform Alberta's personal health record (PHR) design. These two examples demonstrate how to design a descriptive study applying a range of data collection and analysis methods to achieve the evaluation objectives.

11.4.1 United Kingdom HealthSpace Case Study

Between 2007 and 2010, an independent evaluation was commissioned by the U.K. Department of Health to evaluate the implementation and impact of the summary care record (SCR) and HealthSpace programmes (Greenhalgh, Stramer et al., 2010; Greenhalgh, Hinder et al., 2010). SCR was an electronic summary of key health data drawn from a patient's GP-held electronic record

and accessible over a secure Internet connection by authorized healthcare staff. HealthSpace was an Internet-accessible personal organizer onto which people may enter health data and plan health appointments. Through an advanced HealthSpace account, they could gain secure access to their SCR and e-mail their GP using a function called Communicator.

This evaluation undertook a mixed methods approach using a range of data sources and collection methods to "capture as rich a picture of the programme as possible from as many angles as possible" (Greenhalgh, Hinder et al., 2010). The evaluation fieldwork involved seven interrelated empirical studies, including a multilevel case study of HealthSpace covering the policy-making process, implementation by the English National Health Service (NHS) organizations, and experiences of patients and carers. In the case study, evaluators reviewed the national registration statistics on the HealthSpace uptake rate (using the number of basic and advanced HealthSpace accounts created). They also studied the adoption and non-adoption of HealthSpace by 56 patients and carers using observation and interview methods. In addition, they interviewed 160 staff in national and local organizations, and collected 3,000 pages of documents to build a picture of the programme in context. As part of the patient study, ethnographic observation was undertaken by a researcher who shadowed 20 participants for two or three periods of two to five hours each at home and work, and noted information needs as they arose and how these were tackled by the participant. An in-depth picture of HealthSpace conception, design, implementation, utilization (or non-use and abandonment, in most cases) and impact was constructed from this mixed methods approach that included both quantitative uptake statistics and qualitative analysis of the field notes, interview transcripts, documents and communication records.

The case study showed that the HealthSpace personal electronic health record was poorly taken up by people in England, and it was perceived as neither useful nor easy to use. The study also made several recommendations for future development of similar technologies, including the suggestion to conceptualize them as components of a sociotechnical network and to apply user-centred design principles more explicitly. The overall evaluation of the SCR and HealthSpace recognized the scale and complexity of both programmes and observed that "greatest progress appeared to be made when key stakeholders came together in uneasy dialogue, speaking each other's languages imperfectly and trying to understand where others were coming from, even when the hoped-for consensus never materialised" (Greenhalgh, Hinder et al., 2010).

11.4.2 Usability Evaluation to Inform Alberta's PHR Design

The Alberta PHR was a key component in the online consumer health application, the Personal Health Portal (PHP), deployed in the Province of Alberta, Canada. The PHR usability evaluation (Price, Bellwood, & Davies, 2015) was part of the overall PHP benefit evaluation that was embedded into the life cycle of the PHP program throughout the predesign, design and adoption phases. Al-

though using a commercial PHR product, its usability evaluation aimed to assess the early design of the PHR software and to provide constructive feedback and recommendations to the PHR project team in a timely way so as to improve the PHR software prior to its launch.

Between June 2012 and April 2013, a combination of usability inspection (applying heuristic inspection and persona-based inspection methods) and usability testing (with 21 representative end users) was used in Alberta's PHR evaluation. For the persona-based inspection, two patient personas were developed; for each persona, scenarios were developed to illustrate expected use of the PHR. Then in the user testing protocol, participants were asked to "think aloud" while performing two sets of actions: (a) to explore the PHR freely, and (b) to follow specific scenarios matching the expected activities of the targeted end users that covered all key PHR tasks. Findings from the usability inspection and testing were largely consistent and were used to generate several recommendations regarding the PHR information architecture, content and presentation. For instance, the usability inspection identified that the PHR had a deep navigation hierarchy with several layers of screens before patient health data became available. This was also confirmed in usability testing when users sometimes found the module segmentation confusing. Accordingly, the evaluation researchers have recommended revising the structure and organization of the modules with clearer top-level navigation, a combination of content-oriented tabs and user-specific tabs, and a "home" tab providing a clear clinical summary.

Usability evaluation can be conducted at several stages in the development life cycle of eHealth systems to improve the design — from the earliest mockups (ideally starting with paper prototypes), on partially completed systems, or once the system is installed and undergoing maintenance. The Alberta PHR study represents an exemplary case of usability evaluations to inform the development of a government-sponsored PHR project. It demonstrates the feasibility and value of early usability evaluation in eHealth projects for having a good usable system, in this case avoiding usability problems prior to rollout.

11.5 Summary

Descriptive evaluation studies describe the process and impact of the development and implementation of a system. The findings are often contextualized within the implementation environment, such as — for our purposes — the specific healthcare organization. Descriptive evaluations utilize a variety of qualitative and quantitative data collection and analysis methods; and the study design can apply a range of assumptions, from positivist or interpretivist perspectives, to critical theory and critical realism. These studies are used in both formative evaluations and summative evaluations.

References

Benbasat, I., Goldstein, D. K., & Mead, M. (1987). The case research strategy in studies of information-systems. *Management Information Systems Quarterly, 11*(3), 369–386. doi: 10.2307/248684

Chiasson, M., Germonprez, M., & Mathiassen, L. (2009). Pluralist action research: a review of the information systems literature. *Information Systems Journal, 19*(1), 31–54. doi: 10.1111/j.1365-2575.2008.00297.x

Cooper, A. (2004). *The inmates are running the asylum: Why high-tech products drive us crazy and how to restore the sanity*. Carmel, CA: Sams Publishing.

Creswell, J. W., Klassen, A. C., Plano Clark, V. L., & Smith, K. C. (2011, August). *Best practices for mixed methods research in the health sciences.* Bethesda, MD: Office of Behavioral and Social Sciences Research, National Institutes of Health. Retrieved from http://obssr.od.nih.gov/mixed_methods_research

Dubé, L., & Paré, G. (2003). Rigor in information systems positivist case research: Current practices, trends, and recommendations. *Management Information Systems Quarterly, 27*(4), 597–635.

Friedman, C., & Wyatt, J. (1997). *Evaluation methods in medical informatics.* New York: Springer-Verlag.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research.* Chicago: Aldine Pub. Co.

Greenhalgh, T., Hinder, S., Stramer, K., Bratan, T., & Russell, J. (2010). Adoption, non-adoption, and abandonment of a personal electronic health record: case study of HealthSpace. *British Medical Journal, 341*(7782), c5814. doi: 10.1136/bmj.c5814.

Greenhalgh, T., Stramer, K., Bratan, T., Byrne, E., Russell, J., Hinder, S., & Potts, H. (2010, May). *The devil's in the detail. Final report of the independent evaluation of the Summary Care Record and HealthSpace programmes.* London: University College London. Retrieved from https://www.ucl.ac.uk/news/scriefullreport.pdf

Harvey, L., & Myers, M. D. (1995). Scholarship and practice: the contribution of ethnographic research methods to bridging the gap. *Information Technology & People, 8*(3), 13–27.

Kushniruk, A. W., & Patel, V. L. (2004). Cognitive and usability engineering methods for the evaluation of clinical information systems. *Journal of Biomedical Informatics, 37*(1), 56–76. doi: 10.1016/j.jbi.2004.01.003

Levers, M.-J. D. (2013). Philosophical paradigms, grounded theory, and perspectives on emergence. *SAGE Open Journals* (October-December). doi: 10.1177/2158244013517243

Matavire, R., & Brown, I. (2013). Profiling grounded theory approaches in information systems research [dagger]. *European Journal of Information Systems, 22*(1), 119–129.

McEvoy, P., & Richards, D. (2003). Critical realism: a way forward for evaluation research in nursing? *Journal of Advanced Nursing, 43*(4), 411–420. doi: 10.1046/j.1365-2648.2003.02730.x

McKibbon, A. (2015). *eHealth evaluation: Introduction to qualitative methods.* Waterloo, ON: National Institutes of Health Informatics, Canada. Retrieved from http://www.nihi.ca/index.php?MenuItemID=415

McNiff, J., & Whitehead, J. (2002). *Action research: Principles and practice* (2nd ed.). London: Routledge.

Myers, M. D. (1997a). ICIS Panel 1995: Judging qualitative research in information systems: Criteria for accepting and rejecting manuscripts. Criteria and conventions used for judging manuscripts in the area of ethnography. Retrieved from http://www.misq.org/skin/frontend/default/misq/MISQD_isworld/iciseth.htm

Myers, M. D. (1997b). Qualitative research in information systems. *Management Information Systems Quarterly, 21*(2), 241–242. doi: 10.2307/249422

Neuman, W. L. (2003). *Social research methods: Qualitative and quantitative approach* (5th ed.). Boston: Pearson Education, Inc.

Nielsen, J., & Molich, R. (1990). *Heuristic evaluation of user interfaces.* Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, April 1 to 5, Seattle, Washington, U.S.A.

Polson, P. G., Lewis, C., Rieman, J., & Wharton, C. (1992). Cognitive walkthroughs: A method for theory-based evaluation of user interfaces. *International Journal of Man-Machine Studies, 36*(5), 741–773. doi: 10.1016/0020-7373(92)90039-N

Preece, J., Rogers, Y., & Sharp, H. (2002). *Interaction design: beyond human-computer interaction.* New York: Wiley.

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., & Carey, T. (1994). *Human-computer interaction.* New York: Addison-Wesley Publishing Company.

Price, M., Bellwood, P., & Davies, I. (2015). Using usability evaluation to inform Alberta's personal health record design. *Studies in Health Technology and Informatics, 208*, 314–318.

Reason, P., & Bradbury, H. (Eds.). (2001). *Handbook of action research: Participative inquiry and practice* (1st ed.). London: SAGE Publications.

Rhine, J. (2008, Wednesday, July 23). The Grounded Theory Institute: The official site of Dr. Barney Glaser and classic grounded theory. Retrieved from http://www.groundedtheory.com/

Schwandt, T. A. (2007). *The SAGE dictionary of qualitative inquiry* (3rd ed.). Thousand Oaks, CA: SAGE Publications.

van der Meijden, M. J., Tange, H. J., Troost, J., & Hasman, A. (2003). Determinants of success of inpatient clinical information systems: A literature review. *Journal of American Medical Informatics Association, 10*(3), 235–243. doi: 10.1197/jamia.M1094

Walsham, G. (2006). Doing interpretive research. *European Journal of Information Systems, 15*(3), 320–330.

Wisdom, J. P., Cavaleri, M. A., Onwuegbuzie, A. J., & Green, C. A. (2012). Methodological reporting in qualitative, quantitative, and mixed methods health services research articles. *Health Services Research, 47*(2), 721–745. doi: 10.1111/j.1475-6773.2011.01344.x

Yin, R. K. (2011). *Case study research: Design and methods* (Vol. 5). Thousand Oaks, CA: SAGE Publications.

Chapter 12

Methods for Correlational Studies

Francis Lau

12.1 Introduction

Correlational studies aim to find out if there are differences in the characteristics of a population depending on whether or not its subjects have been exposed to an event of interest in the naturalistic setting. In eHealth, correlational studies are often used to determine whether the use of an eHealth system is associated with a particular set of user characteristics and/or quality of care patterns (Friedman & Wyatt, 2006). An example is a computerized provider order entry (CPOE) study to differentiate the background, usage and performance between clinical users and non-users of the CPOE system after its implementation in a hospital.

Correlational studies are different from comparative studies in that the evaluator does not control the allocation of subjects into comparison groups or assignment of the intervention to specific groups. Instead, the evaluator defines a set of variables including an outcome of interest then tests for hypothesized relations among these variables. The outcome is known as the dependent variable and the variables being tested for association are the independent variables. Correlational studies are similar to comparative studies in that they take on an objectivist view where the variables can be defined, measured and analyzed for the presence of hypothesized relations. As such, correlational studies face the same challenges as comparative studies in terms of their internal and external validity. Of particular importance are the issues of design choices, selection bias, confounders, and reporting consistency.

In this chapter we describe the basic types of correlational studies seen in the eHealth literature and their methodological considerations. Also included are three case examples to show how these studies are done.

12.2 Types of Correlational Studies

Correlational studies, better known as observational studies in epidemiology, are used to examine event exposure, disease prevalence and risk factors in a population (Elwood, 2007). In eHealth, the exposure typically refers to the use of an eHealth system by a population of subjects in a given setting. These subjects may be patients, providers or organizations identified through a set of variables that are thought to differ in their measured values depending on whether or not the subjects were "exposed" to the eHealth system.

There are three basic types of correlational studies that are used in eHealth evaluation: cohort, cross-sectional, and case-control studies (Vandenbroucke et al., 2014). These are described below.

- *Cohort studies* – A sample of subjects is observed over time where those exposed and not exposed to the eHealth system are compared for differences in one or more predefined outcomes, such as adverse event rates. Cohort studies may be prospective in nature where subjects are followed for a time period into the future or retrospective for a period into the past. The comparisons are typically made at the beginning of the study as baseline measures, then repeated over time at predefined intervals for differences and trends. Some cohort studies involve only a single group of subjects. Their focus is to describe the characteristics of subjects based on a set of variables, such as the pattern of EHR use by providers and their quality of care in an organization over a given time period.

- *Cross-sectional studies* – These are considered a type of cohort study where only one comparison is made between exposed and unexposed subjects. They provide a snapshot of the outcome and the associated characteristics of the cohort at a specific point in time.

- *Case-control studies* – Subjects in a sample that are exposed to the eHealth system are matched with those not exposed but otherwise similar in composition, then compared for differences in some predefined outcomes. Case-control studies are retrospective in nature where subjects already exposed to the event are selected then matched with unexposed subjects, using historical cases to ensure they have similar characteristics.

A cross-sectional survey is a type of cross-sectional study where the data source is drawn from postal questionnaires and interviews. This topic will be covered in the chapter on methods for survey studies.

12.3 Methodological Considerations

While correlational studies are considered less rigorous than RCTs, they are the preferred designs when it is neither feasible nor ethical to conduct experimental trials. Key methodological issues arise in terms of: (a) design options, (b) biases and confounders, (c) controlling for confounding effects, (d) adherence to good practices, and (e) reporting consistency. These issues are discussed below.

12.3.1 Design Options

There are growing populations with multiple chronic conditions and healthcare interventions. They have made it difficult to design RCTs with sufficient sample size and long-term follow-up to account for all the variability this phenomenon entails. Also RCTs are intended to test the efficacy of an intervention in a restricted sample of subjects under ideal settings. They have limited generalizability to the population at large in routine settings (Fleurence, Naci, & Jansen, 2010). As such, correlational studies, especially those involving the use of routinely collected EHR data from the general population, have become viable alternatives to RCTs. There are advantages and disadvantages to each of the three design options presented above. They are listed below.

- *Cohort studies* – These studies typically follow the cohorts over time, which allow one to examine causal relationships between exposure and one or more outcomes. They also allow one to measure change in exposure and outcomes over time. However, these studies can be costly and time-consuming to conduct if the outcomes are rare or occur in the future. With prospective cohorts they can be prone to dropout. With retrospective cohorts accurate historical records are required which may not be available or complete (Levin, 2003a).

- *Case-control studies* – These studies are suited to examine infrequent or rare outcomes since they are selected at the outset to ensure sufficient cases. Yet the selection of exposed and matching cases can be problematic, as not all relevant characteristics are known. Moreover, the cases may not be representative of the population of interest. The focus on exposed cases that occur infrequently may overestimate their risks (Levin, 2003b).

- *Cross-sectional studies* – These studies are easier and quicker to conduct than others as they involve a one-time effort over a short period using a sample from the population of interest. They can be used to generate hypotheses and examine multiple outcomes and characteristics at the same time with no loss to follow-up. On the other hand, these studies only give a snapshot of the situation at one time point, making it difficult for causal inference of the ex-

posure and outcomes. The results might be different had another time period been chosen (Levin, 2006).

12.3.2 Biases and Confounders

Shamliyan, Kane, and Dickinson (2010) conducted a systematic review on tools used to assess the quality of observational studies. Despite the large number of quality scales and checklists found in the literature, they concluded that the universal concerns are in the areas of selection bias, confounding, and misclassification. These concerns, also mentioned by Vandenbroucke and colleagues (2014) in their reporting guidelines for observational studies, are summarized below.

- *Selection bias* – When subjects are selected through their exposure to the event rather than by random or concealed allocation, there is a risk that the subjects are not comparable due to the presence of systematic differences in their baseline characteristics. For example, a correlational study that examines the association between EHR use and quality of care may have younger providers with more computer savvy in the exposed group because they use EHR more and with more facility than those in the unexposed group. It is also possible to have sicker patients in the exposed group since they require more frequent EHR use than unexposed patients who may be healthier and have less need for the EHR. This is sometimes referred to as response bias, where the characteristics of subjects agreed to be in the study are different from those who declined to take part.

- *Confounding* – Extraneous factors that influence the outcome but are also associated with the exposure are said to have a confounding effect. One such type is confounding by indication where sicker patients are both more likely to receive treatments and also more likely to have adverse outcomes. For example, a study of CDS alerts and adverse drug events may find a positive but spurious association due to the inclusion of sicker patients with multiple conditions and medications, which increases their chance of adverse events regardless of CDS alerts.

- *Misclassification* – When there are systematic differences in the completeness or accuracy of the data recorded on the subjects, there is a risk of misclassification in their exposures or outcomes. This is also known as information or detection bias. An example is where sicker patients may have more complete EHR data because they received more tests, treatments and outcome tracking than those who are healthier and require less attention. As such, the exposure and outcomes of sicker patients may be overestimated.

It is important to note that bias and confounding are not synonymous. Bias is caused by finding the wrong association from flawed information or subject selection. Confounding is factually correct with respect to the relationship found, but is incorrect in its interpretation due to an extraneous factor that is associated with both the exposure and outcome.

12.3.3 Controlling for Confounding Effects

There are three common methods to control for confounding effects. These are by matching, stratification, and modelling. They are described below (Higgins & Green, 2011).

- *Matching* – The selection of subjects with similar characteristics so that they are comparable; the matching can be done at the individual subject level where each exposed subject is matched with one or more unexposed subjects as controls. It can also be done at the group level with equal numbers of exposed and unexposed subjects. Another way to match subjects is by propensity score, that is, a measure derived from a set of characteristics in the subjects. An example is the retrospective cohort study by Zhou, Leith, Li, and Tom (2015) to examine the association between caregiver PHR use and healthcare utilization by pediatric patients. In that study, a propensity score-matching algorithm was used to match PHR-registered children to non-registered children. The matching model used registration as the outcome variable and all child and caregiver characteristics as the independent variables.

- *Stratification* – Subjects are categorized into subgroups based on a set of characteristics such as age and sex then analyzed for the effect within each subgroup. An example is the retrospective cohort study by Staes et al. (2008), examining the impact of computerized alerts on the quality of outpatient lab monitoring for transplant patients. In that study, the before/after comparison of the timeliness of reporting and clinician responses was stratified by the type of test (creatinine, cyclosporine A, and tacrolimus) and report source (hospital laboratory or other labs).

- *Modelling* – The use of statistical models to compute adjusted effects while accounting for relevant characteristics such as age and sex differences among subjects. An example is the retrospective cohort study by Beck and colleagues (2012) to compare documentation consistency and care plan improvement before and after the implementation of an electronic asthma-specific history and phys-

ical template. In that study, before/after group characteristics were compared for differences using *t*-tests for continuous variables and χ^2 statistics for categorical variables. Logistic regression was used to adjust for group differences in age, gender, insurance, albuterol use at admission, and previous hospitalization.

12.3.4 Adherence to Good Practices in Prospective Observational Studies

The ISPOR[1] Good Research Practices Task Force published a set of recommendations in designing, conducting and reporting prospective observational studies for comparative effectiveness research (Berger et al., 2012) that are relevant to eHealth evaluation. Their key recommendations are listed below.

- Key policy questions should be defined to allow inferences to be drawn.
- Hypothesis testing protocol design to include the hypothesis/questions, treatment groups and outcomes, measured and unmeasured confounders, primary analyses, and required sample size.
- Rationale for prospective observational study design over others (e.g., RCT) is based on question, feasibility, intervention characteristics and ability to answer the question versus cost and timeliness.
- Study design choice is able to address potential biases and confounders through the use of inception cohorts, multiple comparator groups, matching designs and unaffected outcomes.
- Explanation of study design and analytic choices is transparent.
- Study execution is carried out in ways that ensure relevance and reasonable follow-up is not different from the usual practice.
- Study registration takes place on publicly available sites prior to its initiation.

12.3.5 The Need for Reporting Consistency

Vandenbroucke et al. (2014) published an expanded version of the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement to improve the reporting of observational studies that can be applied in eHealth evaluation. It is made up of 22 items, of which 18 are com-

1 ISPOR – International Society for Pharmacoeconomics and Outcomes Research

mon to cohort, case-control and cross-sectional studies, with four being specific to each of the three designs. The 22 reporting items are listed below (for details refer to the cited reference).

- *Title and abstract* – one item that covers the type of design used, and a summary of what was done and found.
- *Introduction* – two items on study background/rationale, objectives and/or hypotheses.
- *Methods* – nine items on design, setting, participants, variables, data sources/measurement, bias, study size, quantitative variables and statistical methods used.
- *Results* – five items on participants, descriptive, outcome data, main results and other analyses.
- *Discussion* – four items on key results, limitations, interpretation and generalizability.
- *Other information* – one item on funding source.

The four items specific to study design relate to the reporting of participants, statistical methods, descriptive results and outcome data. They are briefly described below for the three types of designs.

- *Cohort studies* – Participant eligibility criteria and sources, methods of selection, follow-up and handling dropouts, description of follow-up time and duration, and number of outcome events or summary measures over time. For matched studies include matching criteria and number of exposed and unexposed subjects.
- *Cross-sectional studies* – Participant eligibility criteria, sources and methods of selection, analytical methods accounting for sampling strategy as needed, and number of outcome events or summary measures.
- *Case-control studies* – Participant eligibility criteria, sources and methods of case/control selection with rationale for choices, methods of matching cases/controls, and number of exposures by category or summary measures of exposures. For matched studies include matching criteria and number of controls per case.

12.4 Case Examples

12.4.1 Cohort Study of Automated Immunosuppressive Care

Park and colleagues (2010) conducted a retrospective cohort study to examine the association between the use of a CDS (clinical decision support) system in post-liver transplant immunosuppressive care and the rates of rejection episode and drug toxicity. The study is summarized below.

- *Setting* – A liver transplant program in the United States that had implemented an automated CDS system to manage immunosuppressive therapy for its post-liver transplant recipients after discharge. The system consolidated all clinical information to expedite immunosuppressive review, ordering, and follow-up with recipients. Prior to automation, a paper charting system was used that involved manually tracking lab tests, transcribing results into a paper spreadsheet, finding physicians to review results and orders, and contacting recipients to notify them of changes.

- *Subjects* – The study population included recipients of liver transplants between 2004 and 2008 who received outpatient immunosuppressive therapy that included tacrolimus medications.

- *Design* – A retrospective cohort study with a before/after design to compare recipients managed by the paper charting system against those managed by the CDS system for up to one year after discharge.

- *Measures* – The outcome variables were the percentages of recipients with at least one rejection and/or tacrolimus toxicity episode during the one-year follow-up period. The independent variables included recipient, intraoperative, donor and postoperative characteristics, and use of paper charting or CDS. Examples of recipient variables were age, gender, body mass index, presence of diabetes and hypertension, and pre-transplant lab results. Examples of intraoperative data were blood type match, type of transplant and volume of blood transfused. Examples of donor data included percentage of fat in the liver. Examples of post-transplantation data included the type of immunosuppressive induction therapy and the management method.

- *Analysis* – Mean, standard deviation and *t*-tests were computed for continuous variables after checking for normal distribution. Percentages and Fisher's exact test were computed for categorical variables. Autoregressive integrated moving average analysis was

done to determine change in outcomes over time. Logistic regression with variables thought to be clinically relevant was used to identify significant univariable and multivariable factors associated with the outcomes. *P* values of less than 0.05 were considered significant.

- *Findings* – Overall, the CDS system was associated with significantly fewer episodes of rejection and tacrolimus toxicity. The integrated moving average analysis showed a significant decrease in outcome rates after the CDS system was implemented compared with paper charting. Multivariable analysis showed the CDS system had lower odds of a rejection episode than paper charting (OR 0.20; $p < 0.01$) and lower odds of tacrolimus toxicity (OR 0.5; $p < 0.01$). Other significant non-system related factors included the use of specific drugs, the percentage of fat in the donor liver and the volume of packed red cells transfused.

12.4.2 Cross-sectional Analysis of EHR Documentation and Care Quality

Linder, Schnipper, and Middleton (2012) conducted a cross-sectional study to examine the association between the type of EHR documentation used by physicians and the quality of care provided. The study is summarized below.

- *Setting* – An integrated primary care practice-based research network affiliated with an academic centre in the United States. The network uses an in-house EHR system with decision support for preventive services, chronic care management, and medication monitoring and alerts. The EHR data include problem and medication lists, coded allergies and lab tests.

- *Subjects* – Physicians and patients from 10 primary care practices that were part of an RCT to examine the use of a decision support tool to manage patients with coronary artery disease and diabetes (CAD/DM). Eligible patients were those with CAD/DM in their EHR problem list prior to the RCT start date.

- *Design* – A nine-month retrospective cross-sectional analysis of EHR data collected from the RCT. Three physician documentation styles were defined based on 188,554 visit notes in the EHR: (a) dictation, (b) structured documentation, and (c) free text note. Physicians were divided into three groups based on their predominant style defined as more than 25% of their notes composed by a given method.

- *Measures* – The outcome variables were 15 EHR-based CAD/DM quality measures assessed 30 days after primary care visits. They covered quality of documentation, medication use, lab testing, physiologic measures, and vaccinations. Measures collected prior to the day of visit were eligible and considered fulfilled with the presence of coded EHR data on vital signs, medications, allergies, problem lists, lab tests, and vaccinations. Independent variables on physicians and patients were included as covariates. For physicians, they included age, gender, training level, proportion of CAD/DM patients in their panel, total patient visits, and self-reported experience with the EHR. For patients, they included sociodemographic factors, the number of clinic visits and hospitalizations, the number of problems and medications in the EHR, and whether their physician was in the intervention group.

- *Analysis* – Baseline characteristics of physicians and patients were compared using descriptive statistics. Continuous variables were compared using ANOVA. For categorical variables, Fisher's exact test was used for physician variables and χ^2 test for patient variables. Multivariate logistic regression models were used for each quality measure to adjust for patient and physician clustering and potential confounders. Bonferroni procedure was used to account for multiple comparisons for the 15 quality measures.

- *Findings* – During the study period, 234 physicians documented 18,569 visits from 7,000 CAD/DM patients. Of these physicians, 146 (62%) typed free-text notes, 68 (25%) used structured documentation, and 20 (9%) dictated notes. After adjusting for cluster effect, physicians who dictated their notes had the worst quality of care in all 15 measures. In particular, physicians who dictated notes were significantly worse in three of 15 measures (antiplatelet medication, tobacco use, diabetic eye exam); physicians who used structured documentation were better in three measures (blood pressure, body mass, diabetic foot exam); and those who used free-text were better in one measure (influenza vaccination). In summary, physicians who dictated notes had worse quality of care than those with structured documentation.

12.4.3 Case-control Comparison of Internet Portal Use

Nielsen, Halamka, and Kinkel (2012) conducted a case-control study to evaluate whether there was an association between active Internet patient portal use by Multiple Sclerosis (MS) patients and medical resource utilization. Patient predictors and barriers to portal use were also identified. The study is summarized below.

- *Setting* – An academic MS centre in the United States with an in-house Internet patient portal site that was accessed by MS patients to schedule clinic appointments, request prescription refills and referrals, view test results, upload personal health information, and communicate with providers via secure e-mails.

- *Subjects* – 240 adult MS patients actively followed during 2008 and 2009 were randomly selected from the EHR; 120 of these patients had submitted at least one message during that period and were defined as portal users. Another 120 patients who did not enrol in the portal or send any message were selected as non-users for comparison.

- *Design* – A retrospective case-control study facilitated through a chart review comparing portal users against non-users from the same period. Patient demographic and clinical information was extracted from the EHR, while portal usage, including feature access type and frequency and e-mail message content, were provided by IT staff.

- *Measures* – Patient variables included age, gender, race, insurance type, employment status, number of medical problems, disease duration, psychiatric history, number of medications, and physical disability scores. Provider variables included prescription type and frequency. Portal usage variables included feature access type and frequency for test results, appointments, prescription requests and logins, and categorized messaging contents.

- *Analysis* – Comparison of patient demographic, clinical and medical resource utilization data from users and non-users were made using descriptive statistics, Wilcoxon rank sum test, Fisher's exact test and χ^2 test. Multivariate logistic regression was used to identify patient predictors and barriers to portal use. Provider prescribing habits against patient's psychiatric history and portal use were examined by two-way analysis of variance. All statistical tests used p value of 0.05 with no adjustment made for multiple comparisons. A logistic multivariate regression model was created to predict portal use based on patient demographics, clinical condition, socio-economic status, and physical disability metrics.

- *Findings* – Portal users were mostly young professionals with little physical disability. The most frequently used feature was secure patient-provider messaging, often for medication requests or refills, and self-reported side effects. Predictors and barriers of portal

use were the number of medications prescribed (OR 1.69, $p < 0.0001$), Caucasian ethnicity (OR 5.04, $p = 0.007$), arm and hand disability (OR 0.23, $p = 0.01$), and impaired vision (OR 0.31, $p = 0.01$). For medical resource utilization, portal users had more frequent clinic visits, medication use and prescriptions from centre staff providers. Patients with a history of psychiatric disease were prescribed more MS medications than those without any history ($p < 0.0001$). In summary, MS patients used the Internet more than the general population, but physical disability limited their access and need to be addressed.

12.4.4 Limitations

A general limitation of a correlational study is that it can determine association between exposure and outcomes but cannot predict causation. The more specific limitations of the three case examples cited by the authors are listed below.

- *Automated immunosuppressive care* – Baseline differences existed between groups with unknown effects; possible other unmeasured confounders; possible Hawthorne effects from focus on immunosuppressive care.

- EHR *documentation and care quality* – Small sample size; only three documentation styles were considered (e.g., scribe and voice recognition software were excluded) and unsure if they were stable during study period; quality measures specific to CAD/DM conditions only; complex methods of adjusting for clustering and confounding that did not account for unmeasured confounders; the level of physician training (e.g., attending versus residents) not adjusted.

- *Internet portal use* – Small sample size not representative of the study population; referral centre site could over-represent complex patients requiring advanced care; all patients had health insurance.

12.5 Summary

In this chapter we described cohort, case-control and cross-sectional studies as three types of correlational studies used in eHealth evaluation. The methodological issues addressed include bias and confounding, controlling for confounders, adherence to good practices and consistency in reporting. Three case examples were included to show how eHealth correlational studies are done.

References

Beck, A. F., Sauers, H. S., Kahn, R. S., Yau, C., Weiser, J., & Simmons, J. M. (2012). Improved documentation and care planning with an asthma-specific history and physical. *Hospital Pediatrics, 2*(4), 194–201.

Berger, M. L., Dreyer, N., Anderson, F., Towse, A., Sedrakyan, A., & Normand, S. L. (2012). Prospective observational studies to address comparative effectiveness: The ISPOR good research practices task force report. *Value in Health, 15*(2), 217–230. Retrieved from http://www.sciencedirect.com/science/article/pii/S1098301512000071

Elwood, M. (2007). *Critical appraisal of epidemiological studies and clinical studies* (3rd ed.). Oxford: Oxford University Press.

Fleurence, R. L., Naci, H., & Jansen, J. P. (2010). The critical role of observational evidence in comparative effectiveness research. *Health Affairs, 29*(10), 1826–1833.

Friedman, C. P., & Wyatt, J. C. (2006). *Evaluation methods in biomedical informatics* (2nd ed.). New York: Springer Science + Business Media, Inc.

Higgins, J. P. T., & Green, S. (Eds.). (2011). *Cochrane handbook for systematic reviews of interventions* (Version 5.1.0, updated March 2011). London: The Cochrane Collaboration. Retrieved from http://handbook.cochrane.org/

Levin, K. A. (2003a). Study design IV: Cohort studies. *Evidence-based Dentistry, 7*, 51–52.

Levin, K. A. (2003b). Study design V: Case-control studies. *Evidence-based Dentistry, 7*, 83–84.

Levin, K. A. (2006). Study design *III*: Cross-sectional studies. *Evidence-based Dentistry, 7*, 24–25.

Linder, J. A., Schnipper, J. L., & Middleton, B. (2012). Method of electronic health record documentation and quality of primary care. *Journal of the American Medical Informatics Association, 19*(6), 1019–1024.

Nielsen, A. S., Halamka, J. D., & Kinkel, R. P. (2012). Internet portal use in an academic multiple sclerosis center. *Journal of the American Medical Informatics Association, 19*(1), 128–133.

Park, E. S., Peccoud, M. R., Wicks, K. A., Halldorson, J. B., Carithers Jr., R. L., Reyes, J. D., & Perkins, J. D. (2010). Use of an automated clinical management system improves outpatient immunosuppressive care following liver transplantation. *Journal of the American Medical Informatics Association, 17*(4), 396–402.

Shamliyan, T., Kane, R. L., & Dickinson, S. (2010). A systematic review of tools used to assess the quality of observational studies that examine incidence or prevalence and risk factors for diseases. *Journal of Clinical Epidemiology, 63*(10), 1061–1070.

Staes, C. J., Evans, R. S., Rocha, B. H. S. C., Sorensen, J. B., Huff, S. M., Arata, J., & Narus, S. P. (2008). Computerized alerts improve outpatient laboratory monitoring of transplant patients. *Journal of the American Medical Informatics Association, 15*(3), 324–332.

Vandenbroucke, J. P., von Elm, E., Altman, D. G., Gotzsche, P. C., Mulrow, C. D., Pocock, S. J., … Egger, M., for the STROBE Initiative. (2014). Strengthening the reporting of observational studies in epidemiology (STROBE): explanation and elaboration. *International Journal of Surgery, 12*(12), 1500–1524. Retrieved from http://www.sciencedirect.com/science/article/pii/s174391911400212x

Zhou, Y. Y., Leith, W. M., Li, H., & Tom, J. O. (2015). Personal health record use for children and health care utilization: propensity score-matched cohort analysis. *Journal of the American Medical Informatics Association, 22*(4), 748–754.

Chapter 13

Methods for Survey Studies

Francis Lau

13.1 Introduction

The survey is a popular means of gauging people's opinion of a particular topic, such as their perception or reported use of an eHealth system. Yet surveying as a scientific approach is often misconstrued. And while a survey seems easy to conduct, ensuring that it is of high quality is much more difficult to achieve. Often the terms "survey" and "questionnaire" are used interchangeably as if they are the same. But strictly speaking, the survey is a research approach where subjective opinions are collected from a sample of subjects and analyzed for some aspects of the study population that they represent. On the other hand, a questionnaire is one of the data collection methods used in the survey approach, where subjects are asked to respond to a predefined set of questions.

The eHealth literature is replete with survey studies conducted in different health settings on a variety of topics, for example the perceived satisfaction of EHR systems by ophthalmologists in the United States (Chiang et al., 2008), and the reported impact of EMR adoption in primary care in a Canadian province (Paré et al., 2013). The quality of eHealth survey studies can be highly variable depending on how they are designed, conducted, analyzed and reported. It is important to point out there are different types of survey studies that range in nature from the exploratory to the predictive, involving one or more groups of subjects and an eHealth system over a given time period. There are also various published guidelines on how survey studies should be designed, reported and appraised. Increasingly, survey studies are used by health organizations to learn about provider, patient and public perceptions toward eHealth systems. As a consequence, the types of survey studies and their methodological considerations should be of great interest to those involved with eHealth evaluation.

This chapter describes the types of survey studies used in eHealth evaluation and their methodological considerations. Also included are three case examples to show how these studies are done.

13.2 Types of Survey Studies

There are different types of survey study designs depending on the intended purpose and approach taken. Within a given type of survey design, there are different design options with respect to the time period, respondent group, variable choice, data collection and analytical method involved. These design features are described below (Williamson & Johanson, 2013).

13.2.1 The Purpose of Surveys

There are three broad types of survey studies reported in the eHealth literature: exploratory, descriptive, and explanatory surveys. They are described below.

- *Exploratory Surveys* – These studies are used to investigate and understand a particular issue or topic area without predetermined notions of the expected responses. The design is mostly qualitative in nature, seeking input from respondents with open-ended questions focused on why and/or how they perceive certain aspects of an eHealth system. An example is the survey by Wells, Rozenblum, Park, Dunn, and Bates (2014) to identify organizational strategies that promote provider and patient uptake of PHRs.

- *Descriptive Surveys* – These studies are used to describe the perception of respondents and the association of their characteristics with an eHealth system. Perception can be the attitudes, behaviours and reported interactions of respondents with the eHealth system. Association refers to an observed correlation between certain respondent characteristics and the system, such as prior eHealth experience. The design is mostly quantitative and involves the use of descriptive statistics such as frequency distributions of Likert scale responses from participants. An example is the survey on change in end user satisfaction with CPOE over time in intensive care (Hoonakker et al., 2013).

- *Explanatory Surveys* – These studies are used to explain or predict one or more hypothesized relationships between some respondent characteristics and the eHealth system. The design is quantitative, involving the use of inferential statistics such as regression and factor analysis to quantify the extent to which certain respondent characteristics lead to or are associated with specific outcomes. An example is the survey to model certain res-

idential care facility characteristics as predictors of EHR use (Holup, Dobbs, Meng, & Hyer, 2013).

13.2.2 Survey Design Options

Within the three broad types of survey studies one can further distinguish their design by time period, respondent group, variable choice, data collection and analytical method. These survey design options are described below.

- *Time Period* – Surveys can take on a cross-sectional or longitudinal design based on the time period involved. In cross-sectional design the survey takes place at one point in time giving a snapshot of the participant responses. In longitudinal design the survey is repeated two or more times within a specified period in order to detect changes in participant responses over time.

- *Respondent Group* – Surveys can involve a single or multiple cohorts of respondents. With multiple cohorts they are typically grouped by some characteristics for comparison such as age, sex, or eHealth use status (e.g., users versus non-users of EMR).

- *Variable Choice* – In quantitative surveys one needs to define the dependent and independent variables being studied. A dependent variable refers to the perceived outcome that is measured, whereas an independent variable refers to a respondent characteristic that may influence the outcome (such as age). Typically the variables are defined using a scale that can be nominal, ordinal, interval, or ratio in nature (Layman & Watzlaf, 2009). In a nominal scale, a value is assigned to each response such as 1 or F for female and 2 or M for male. In an ordinal scale, the response can be rank ordered such as user satisfaction that starts from 1 for very unsatisfied to 4 for very satisfied. Interval and ratio scales have numerical meaning where the distance between two responses relate to the numerical values assigned. Ratio is different from interval in that it has a natural zero point. Two examples are weight as a ratio scale and temperature as an interval scale.

- *Data Collection* – Surveys can be conducted by questionnaire or by interview with structured, semi-structured or non-structured questions. Questionnaires can be administered by postal mail, telephone, e-mail, or through a website. Interviews can be conducted in-person or by phone individually or in groups. Pretesting or pilot testing of the instrument should be done with a small number of individuals to ensure its content, flow and instructions are clear,

consistent, appropriate and easy to follow. Usually there are one or more follow-up reminders sent to increase the response rate.

- *Analytical Method* – Survey responses are analyzed in different ways depending on the type of data collected. For textual data such qualitative analyses as content or thematic analysis can be used. Content analysis focuses on classifying words and phrases within the texts into categories based on some initial coding scheme and frequency counts. Thematic analysis focuses on identifying concepts, relationships and patterns from texts as themes. For numeric data, quantitative analysis such as descriptive and inferential statistics can be used. Descriptive statistics involves the use of such measures as mean, range, standard deviation and frequency to summarize the distribution of numeric data. Inferential statistics involve the use of a random sample of data from the study population to make inferences about that population. The inferences are made with parametric and non-parametric tests and multivariate methods. Pearson correlation, *t*-test and analysis of variance are examples of parametric tests. Sign test, Mann-Witney U test and χ^2 are examples of non-parametric tests. Multiple regression, multivariate analysis of variance, and factor analysis are examples of multivariate methods (Forza, 2002).

13.3 Methodological Considerations

The quality of survey studies is dependent on a number of design parameters. These include population and sample, survey instrument, sources of bias, and adherence to reporting standards. These considerations are described below (Williamson & Johanson, 2013).

13.3.1 Population and Sample

For practical reasons, survey studies are often done on a sample of individuals rather than the entire population. Sampling frame refers to the population of interest from which a representative sample is drawn for the study. The two common strategies used to select the study sample are probability and non-probability sampling. These are described below.

- *Probability sampling* – This is used in descriptive and explanatory surveys where the sample selected is based on the statistical probability of each individual being included under the assumption of normal distribution. They include such methods as simple random, systematic, stratified, and cluster sampling. The desired confidence level and margin of error are used to determine the required sample size. For example, in a population of 250,000 at

95% confidence level and a ±5% margin of error, a sample of 384 individuals is needed (Research Advisors, n.d.).

- *Non-probability sampling* – This is used in exploratory surveys where individuals with specific characteristics that can help understand the topic being investigated are selected as the sample. They include such non-statistical methods as convenience, snowball, quota, and purposeful sampling. For example, to study the effects of the Internet on patients with chronic conditions one can employ purposeful sampling where only individuals known to have a chronic disease and access to the Internet are selected for inclusion.

13.3.2 Survey Instrument

The survey instrument is the tool used to collect data from respondents on the topic being investigated. Ideally one should demonstrate that the survey instrument chosen is both valid and reliable for use in the study. Validity refers to whether the items (i.e., predefined questions and responses) in the instrument are accurate in what they intend to measure. Reliability refers to the extent to which the data collected are reproducible when repeated on the same or similar groups of respondents. These two constructs are elaborated below.

- *Validity* – The four types of validity are known as face, content, criterion, and construct validity. Face and content validity are qualitative assessments of the survey instrument for its clarity, comprehensibility and appropriateness. While face validity is typically assessed informally by non-experts, content validity is done formally by experts in the subject matter under study. Criterion and construct validity are quantitative assessments where the instrument is measured against other schemes. In criterion validity the instrument is compared with another reputable test on the same respondents, or against actual future outcomes for the survey's predictive ability. In construct validity the instrument is compared with the theoretical concepts that the instrument purports to represent to see how well the two align with each other.

- *Reliability* – The tests for reliability include test-retest, alternate form and internal consistency. Test-retest reliability correlates results from the same survey instrument administered to the same respondents over two time periods. Alternate form reliability correlates results from different versions of the same instrument on the same or similar individuals. Internal consistency reliability measures how well different items in the same survey that measure the same construct produce similar results.

13.3.3 Sources of Bias

There are four potential sources of bias in survey studies. These are coverage, sampling, non-response, and measurement errors. These potential biases and ways to minimize them are described below.

- *Coverage bias* – This occurs when the sampling frame is not representative of the study population such that certain segments of the population are excluded or under-represented. For instance, the use of the telephone directory to select participants would exclude those with unlisted numbers and mobile devices. To address this error one needs to employ multiple sources to select samples that are more representative of the population. For example, in a telephone survey of consumers on their eHealth attitudes and experience, Ancker, Silver, Miller, and Kaushal (2013) included both landline and cell phone to recruit consumers since young adults, men and minorities tend to be under-represented among those with landlines.

- *Sampling bias* – This occurs when the sample selected for the study is not representative of the population such that the sample values cannot be generalized to the broader population. For example, in their survey of provider satisfaction and reported usage of CPOE, Lee, Teich, Spurr, and Bates (1996) reported different response rates between physicians and nurses, and between medical and surgical staffs, which could affect the generalizability of the results. To avoid sampling bias one should clearly define the target population and sampling frame, employ systematic methods such as stratified or random sampling to select samples, identify the extent and causes of response differences, and adjust the analysis and interpretation accordingly.

- *Non-response bias* – This occurs when individuals who responded to the survey have different attributes than those who did not respond to the survey. For example, in their study to model nurses' acceptance of barcoded medication administration technology, Holden, Brown, Scanlon, and Karsh (2012) acknowledged their less than 50% response rate could have led to non-response bias affecting the accuracy of their prediction model. To address this error one can offer incentives to increase response rate, follow up with non-respondents to find out the reasons for their lack of response, or compare the characteristics of non-respondents with respondents or known external benchmarks for differences (King & He, 2005). Adjustments can then be made when the cause and extent of non-response are known.

- *Measurement bias* – This occurs when there is a difference between the survey results obtained and the true values in the population. One major cause is deficient instrument design due to ambiguous items, unclear instructions, or poor usability. To reduce measurement bias one should apply good survey design practices, adequate pretesting or pilot testing of the instrument, and formal tests for validity and reliability. An example of good Web-based eHealth survey design guidelines is the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) by Eysenbach (2004). The checklist has eight item categories and 31 individual items that can be used by authors to ensure quality design and reporting of their survey studies.

13.3.4 Adherence to Reporting Standards

Currently there are no universally accepted guidelines or standards for reporting survey studies. In the field of management information systems (MIS), Grover, Lee, and Durand (1993) published nine ideal survey methodological attributes for analyzing the quality of MIS survey research. In their review of ideal survey methodological attributes, Ju, Chen, Sun, and Wu (2006) found two frequent problems in survey studies published in three top MIS journals to be the failure to perform statistical tests for non-response errors and not using multiple data collection methods. In healthcare, Kelly, Clark, Brown, and Sitzia (2003) published a checklist of seven key points to be covered when reporting survey studies. They are listed below:

1. Explain the purpose of the study with explicit mention of the research question.
2. Explain why the research is needed and mention previous work to provide context.
3. Provide detail on how study was done that covers: the method and rationale; the instrument with its psychometric properties and references to original development/testing; sample selection and data collection processes.
4. Describe and justify the analytical methods used.
5. Present the results in a concise and factual manner.
6. Interpret and discuss the findings.
7. Present conclusions and recommendations.

In eHealth, Bassi, Lau, and Lesperance (2012) published a review of survey-based studies on the perceived impact of EMR in physician office practices. In the review they used the 9-item assessment tool developed by Grover and colleagues (1993) to appraise the reporting quality of 19 EMR survey studies. Using the 9-item tool a score from 0 to 1 was assigned depending on whether the attribute was present or absent, giving a maximum score of 9. Of the 19 survey studies appraised, the quality scores ranged from 0.5 to 8. Over half of the studies did not include a data collection method, the instrument and its validation with respect to pretesting or pilot testing, and non-respondent testing. Only two studies scored 7 or higher which suggested the reporting of the 19 published EMR survey studies was highly variable. The criteria used in the 9-item tool are listed below.

1 Report the approach used to randomize or select samples.

2 Report a profile of the sample frame.

3 Report characteristics of the respondents.

4 Use a combination of personal, telephone and mail data collection methods.

5 Append the whole or part of the questionnaire in the publication.

6 Adopt a validated instrument or perform a validity or reliability analysis.

7 Perform an instrument pretest.

8 Report on the response rate.

9 Perform a statistical test to justify the loss of data from non-respondents.

13.4 Case Examples

13.4.1 Clinical Informatics Governance for EHR in Nursing

Collins, Alexander, and Moss (2015) conducted an exploratory survey study to understand clinical informatics (CI) governance for nursing and to propose a governance model with recommended roles, partnerships and councils for EHR adoption and optimization. The study is summarized below.

- *Setting* – Integrated healthcare systems in the United States with at least one acute care hospital that had pioneered enterprise-wide EHR implementation projects and had reached the Health Information Management Systems Society (HIMSS) Analytics' EMR Adoption Model (EMRAM) level 6 or greater, or were undergoing enterprise-wide integration, standardization and optimization of existing EHR systems across sites.

- *Population and samples* – Nursing informatics leaders in the role of an executive in an integrated healthcare system who could offer their perspective and lessons learned in their organization's clinical and nursing informatics governance structure and its evolution. The sampling frame was the HIMSS Analytics database that contains detailed information on most U.S. healthcare organizations and their health IT status.

- *Design* – A cross-sectional survey conducted through semi-structured telephone interviews with probing questions.

- *Measures* – The survey had four sections: (a) organizational characteristics; (b) participant characteristics; (c) governance structure; and (d) lessons learned. Questions on governance covered decision-making, committees, collaboration, roles, and facilitators/barriers for success in overall and nursing-specific CI governance.

- *Analysis* – Grounded theory techniques of open, axial and selective coding were used to identify overlapping themes on governance structures and CI roles. Data were collected until thematic saturation in open coding was reached. The CI structures of each organization were drawn, compared and synthesized into a proposed model of CI roles, partnerships and councils for nursing. Initial coding was independently validated among the researchers and group consensus was used in thematic coding to develop the model.

- *Results* – Twelve nursing executives (made up of six chief nursing information officers, four directors of nursing informatics, one chief information officer, and one chief CI officer) were interviewed by phone. For analysis 128 open codes were created and organized into 18 axial coding categories where further selective coding led to four high-level themes for the proposed model. The four themes (with lessons learned included) identified as important are: interprofessional partnerships; defining role-based levels of practice

and competence; integration into existing clinical infrastructure; and governance as an evolving process.

- *Conclusion* – The proposed CI governance model can help understand, shape and standardize roles, competencies and structures in CI practices for nursing, as well as be extended to other domains.

13.4.2 Primary Care EMR Adoption, Use and Impacts

Paré et al. (2013) conducted a descriptive survey study to examine the adoption, use and impacts of primary care EMRs in a Canadian province. The study is summarized below.

- *Setting* – Primary care clinics in the Canadian Province of Quebec that had adopted electronic medical records under the provincial government's EMR adoption incentive and accreditation programs.

- *Population and samples* – The population consisted of family physicians as members of the Quebec Federation of General Practitioners that practice in primary care clinics in the province. The sample had three types of physician respondents that: (a) had not adopted EMR (type-1); (b) had EMR in their clinic but were not using it to support their practice (type-2); or (c) used EMR in their clinic to support their practice (type-3).

- *Design* – A cross-sectional survey in the form of a pretested online questionnaire in English and French accessible via a secure website. E-mail invitations were sent to all members followed by an e-mail reminder. With a sampling frame of 9,166 active family physicians in Quebec, 370 responses would be needed to obtain a representative sample with a 95% confidence interval and a margin of error of ±5%.

- *Measures* – For all three respondent types the measures were clinic and socio-demographic profiles and comments. For type-2 and type-3 respondents, the measures were EMR brand and year of implementation. For type-1 the measures were barriers and intent to adopt EMR. For type-2 the measures were reasons and influencing factors for not using EMR, and intent to use EMR in future. For type-3 the measures were EMR use experience, level and satisfaction, ease of use with advanced EMR features, and individual/organizational impacts associated with EMR use.

- *Analysis* – Descriptive statistics in frequencies, per cent and mean Likert scores were used on selected measures. Key analyses included comparison of frequencies by: socio-demographic and clinic profiles; barrier and adoption intent; EMR feature availability and use; and comparison of mean Likert scores for satisfaction and individual and organizational impacts. Individual impacts included perceived efficiency, quality of care and work satisfaction. Organizational impacts included effects on clinical staff, the clinic's financial position, and clients.

- *Results* – Of 4,845 invited physicians, 780 responded to the survey (16% response rate) that was representative of the population. Just over half of EMR users reported the high cost and complexity in EMR acquisition and deployment as the main barriers. Half of non-users reported their clinics intended to deploy EMR in the next year. EMR users made extensive use of basic EMR features such as clinical notes, lab results and scheduling, but few used clinical decision support and data sharing features. For work organization, EMRs addressed logistical issues with paper systems. For care quality, EMRs improved the quality of clinical notes and safety of care provided but not clinical decision-making. For care continuity, EMRs had poor ability to transfer clinical data among providers.

- *Conclusion* – EMR impacts related to a physician's experience where the perceived benefits were tied to the duration of EMR use. Health organizations should continue to certify EMR products to ensure alignment with the provincial EHR.

13.4.3 Nurses' Acceptance of Barcoded Medication Administration Technology

Holden and colleagues (2012) conducted an explanatory survey study to identify predictors of nurses' acceptance of barcoded medication administration (BCMA) in a U.S. pediatric hospital. The study is summarized below.

- *Setting* – A 236-bed free standing academic pediatric hospital in the midwestern U.S. that had recently adopted BCMA. The hospital also had CPOE, a pharmacy information system and automated medication-dispensing units.

- *Population and Sample* – The population consisted of registered nurses that worked at least 24 hours per week at the hospital. The sample consisted of nurses from three care units that had used BCMA for three or more months.

- *Design* – A cross-sectional paper survey with reminders was conducted to test the hypothesis that BCMA acceptance would be best predicted by a larger set of contextualized variables than the base variables in the Technology Acceptance Model (TAM). A multi-item scales survey instrument, validated in previous studies with several added items, was used. The psychometric properties of the survey scales were pretested with 16 non-study nurses.

- *Measures* – Seven BCMA-related perceptions: ease of use, usefulness for the job, non-specific social influence, training, technical support, usefulness for patient care, and social influence from patients/families. Responses were 7-point scales from not-at-all to a-great-deal. Also tracked were variables for age in five categories, as well as experience measured as job tenure in years and months. Two BCMA acceptance variables: behavioural intention to use and satisfaction.

- *Analysis* – Regression of all subsets of perceptions to identify the best predictors of BCMA acceptance using five goodness-of-fit indicators (i.e., R^2, root mean square error, Mallow's Cp statistics, Akaike information criterion, and Bayesian information criterion). An a priori α criterion of 0.05 was used and 95% confidence intervals were computed around parameter estimates.

- *Results* – Ninety-four of 202 nurses returned a survey (46.5% response rate) but 11 worked less than 24 hours per week and were excluded, leaving a final sample of 83 respondents. Nurses perceived moderate ease of use and low usefulness of BCMA. They perceived moderate or higher social influence to use BCMA, and were moderately positive about BCMA training and technical support. Behavioural intention to use BCMA was high but satisfaction was low. Behavioural intention to use BCMA was best predicted by perceived ease of use, non-specific social influence and usefulness for patient care (56% variance explained). Satisfaction was best predicted by perceived ease of use, usefulness for patient care and social influence from patients/families (76% variances explained).

- *Conclusion* – Predicting BCMA acceptance benefited from using a larger set of perceptions and adapting variables.

13.5 Summary

This chapter introduced three types of surveys, namely exploratory, descriptive and explanatory surveys. The methodological considerations addressed included population and sample, survey instrument, variable choice and reporting standards. Three case examples were also included to show how eHealth survey studies are done.

References

Ancker, J. S., Silver, M., Miller, M. C., & Kaushal, R. (2013). Consumer experience with and attitudes toward health information technology: a nationwide survey. *Journal of American Medical Informatics Association, 20*(1), 152–156.

Bassi, J., Lau, F., & Lesperance, M. (2012). Perceived impact of electronic medical records in physician office practices: A review of survey-based research. *Interactive Journal of Medical Research, 1*(2), e3.1–e3.23.

Chiang, M. F., Boland, M. V., Margolis, J. W., Lum, F., Abramoff, M. D., & Hildbrand, L. (2008). Adoption and perceptions of electronic health record systems by ophthalmologists: An American Academy of Ophthalmology survey. *Ophthalmology*, *115*(9), 1591–1597.

Collins, S. A., Alexander, D., & Moss, J. (2015). Nursing domain of CI governance: recommendations for health IT adoption and optimization. *Journal of American Medical Informatics Association, 22*(3), 697–706.

Eysenbach, G. (2004). Improving the quality of web surveys: the checklist for reporting results of Internet e-surveys (CHERRIES). *Journal of Medical Internet Research, 6*(3), e34.

Forza, C. (2002). Survey research in operations management: a process-based perspective. *International Journal of Operations & Production Management, 22*(2), 152–194.

Grover, V., Lee, C. C., & Durand, D. (1993). Analyzing methodological rigor of MIS survey research from 1980-1989. *Information & Management, 24*(6), 305–317.

Holden, R. J., Brown, R. L., Scanlon, M. C., & Karsh, B.- T. (2012). Modeling nurses' acceptance of bar coded medication administration technology at a pediatric hospital. *Journal of American Medical Informatics Association, 19*(6), 1050–1058.

Holup, A. A., Dobbs, D., Meng, H., & Hyer, K. (2013). Facility characteristics associated with the use of electronic health records in residential care facilities. *Journal of American Medical Informatics Association, 20*(4), 787–791.

Hoonakker, P. L. T., Carayon, P., Brown, R. L., Cartmill, R. S., Wetterneck, T. B., & Walker, J. M. (2013). Changes in end-user satisfaction with computerized provider order entry over time among nurses and providers in intensive care units. *Journal of American Medical Informatics Association, 20*(2), 252–259.

Ju, T. L., Chen, Y. Y., Sun, S. Y., & Wu, C. Y. (2006). Rigor in MIS survey research: in search of ideal survey methodological attributes. *Journal of Computer Information Systems, 47*(2), 112–123.

Kelly, K., Clark, B., Brown, V., & Sitzia, J. (2003). Good practice in the conduct and reporting of survey research. *International Journal for Quality in Health Care, 15*(3), 261–266.

King, W. R., & He, J. (2005). External validity in IS survey research. *Communications of the Association for Information Systems, 16*, 880–894.

Layman, E. J., & Watzlaf, V. J. (2009). *Health informatics research methods: Principles and practice.* Chicago: American Health Information Management Association.

Lee, F., Teich, J. M., Spurr, C. D., & Bates, D. W. (1996). Implementation of physician order entry: user satisfaction and self-reported usage patterns. *Journal of American Medical Informatics Association, 3*(1), 42–55.

Paré, G., de Guinea, A. O., Raymond, L., Poba-Nzaou, P., Trudel, M. C., Marsan, J., & Micheneau, T. (2013, October 31). *Computerization of primary care medical clinics in Quebec: Results from a Survey on EMR adoption, use and impacts.* Montreal: HEC Montreal. Retrieved from https://www.infoway-inforoute.ca/index.php/resources/reports/benefits-evaluation/doc_download/1996-computerization-of-primary-care-medical-clinics-in-quebec-results-from-a-survey-on-emr-adoption-use-and-impacts

Research Advisors. (n.d.). *Sample size table.* Retrieved from http://research advisors.com/tools/SampleSize.htm.

Wells, S., Rozenblum, R., Park, A., Dunn, M., & Bates, D. W. (2014). Organizational strategies for promoting patient and provider uptake of personal records. *Journal of American Medical Informatics Association,* 22(1), 213–222.

Williamson, K., & Johanson, G. (Eds.). (2013). *Research methods: Information, systems and contexts* (1st ed.). Prahan, Victoria, Australia: Tilde University Press.

Chapter 14

Methods for eHealth Economic Evaluation Studies

Francis Lau

14.1 Introduction

A plethora of evaluation methods have been used to examine the economic return of eHealth investments in the literature. These methods offer different ways of determining the "value for money" associated with a given eHealth system that are often based on specific assumptions and needs. However, this diversity has created some ambiguity with respect to when and how one should choose among these methods, ways to maintain the rigour of the process and its reporting, while ensuring relevance of the findings to the organization and stakeholders involved.

This chapter reviews the economic evaluation methods that are used in healthcare, especially those that have been applied in eHealth. It draws on the eHealth Economic Evaluation Framework discussed in chapter 5 by elaborating on the common underlying design, analysis and reporting aspects of the methods presented. In so doing, a better understanding of when and how these methods can be applied in real-world settings is gained. Note that it is beyond the scope of this chapter to describe all known economic evaluation methods in detail. Rather, its focus is to introduce selected methods and the processes involved from the eHealth literature. The Appendix to this chapter presents a glossary of relevant terms with additional reference citations for those interested in greater detail on these methods.

Specifically, this chapter describes the types of eHealth economic evaluation methods reported, the process for identifying, measuring and valuating costs and outcomes and assessing impact, as well as best practice guidance that has

been published. Three brief exemplary cases have been included to illustrate the types of eHealth economic evaluation used and their implication on practice.

14.2 eHealth Economic Evaluation Methods

The basic principle behind economic evaluation is the examination of the costs and outcomes associated with each of the options being considered to determine if they are worth the investment (Drummond, Sculpher, Torrance, O'Brien, & Stoddart, 2005). For eHealth it is the compilation of the resources required to adopt a particular eHealth system option and the consequences derived or expected from the adoption of that system. While there are different types of resources involved they are always expressed in monetary units as the cost. Consequences will depend upon the natural units by which the outcomes are measured and whether they are then aggregated and/or converted into a common unit for comparison.

The type of economic analysis is influenced by how the costs and outcomes are handled. In cost-benefit analysis both the costs and outcomes of the options are expressed and compared in a monetary unit. In cost-effectiveness analysis there is one main outcome that is expressed in its natural unit such as the readmission rate. In cost-consequence analysis there are multiple outcomes reported in their respective units without aggregation such as the readmission rate and hospital length of stay. In cost-minimization analysis the least-cost option is selected assuming all options have equivalent outcomes. In cost-utility analysis the outcome is based on health state preference values such as quality-adjusted life years. Regardless of the type of analysis used, it is important to determine the incremental cost of producing an additional unit of outcome from the options being considered.

Economic evaluation can be done through empirical or modelling studies. In empirical studies, actual cost and outcome data, sometimes supplemented with estimates, are collected as part of a field trial such as a randomized controlled study to determine the impact of an eHealth system. The economic impact is then analyzed and reported alongside the field trial result, which is the clinical impact of the system under consideration. In modelling studies, cost and outcome data are extracted from internal and/or published sources, then analyzed with such decision models as Monte Carlo simulation or logistic regression to project future costs and outcomes over a specified time horizon. Some studies combine both the field trial and modelling approaches by applying the empirical data from the trial to make long-term modelling projections. Regardless of the study design, the evaluation perspective, data sources, time frame, options, and comparison method need to be explicit to ensure the rigour and generalizability of the results.

Two other economic evaluation methods used by healthcare organizations in investment decisions are budget impact analysis and priority setting through program budgeting, and marginal analysis. While these two methods are often

used by key stakeholder groups in investment and disinvestment decisions across a wide range of healthcare services and programs based on overall importance, they are seldom seen in the eHealth literature. Even so, it is important to be aware of these methods and their implications in eHealth.

14.3 Determining Costs, Outcomes and Importance

The process of determining the costs, outcomes and importance of an eHealth system are an integral part of any economic evaluation that needs to be made explicit. The process involves the identification of relevant costs and outcomes, the collection and quantification of costs and outcomes from different data sources, appraisal of their monetary value, and examination of the budgetary impact and overall importance of the eHealth system on the organization and its stakeholder groups (Simoens, 2009). The process is described below.

14.3.1 Identification of Costs and Outcomes

The process of identifying costs and outcomes in eHealth economic evaluation involves the determination of the study perspective, time frame, and types of costs and outcomes to be included (Bassi & Lau, 2013). Perspective is the viewpoint from which the evaluation is being considered, which can be individual, organizational, payer, or societal in nature. Depending on the perspective, certain costs and outcomes may be irrelevant and excluded from the evaluation. For instance, from the perspective of general practitioners who work under a fee-for-service arrangement, the change in their patients' productivity or quality of life may have little relevance to the return on investment of the EMR in their office practice. On the other hand, when the EMR is viewed from a societal perspective, any improvement in the overall population's work productivity and health status is considered a positive return on the investment made.

Since the costs and outcomes associated with the adoption of an eHealth system may accrue differently over time, one has to ensure the time frame chosen for the study is of sufficient duration to capture all of the relevant data involved. For instance, during the implementation of a system there can be decreased staff productivity due to the extra workload and learning required. Similarly, there is often a time delay before the expected change in outcomes can be observed, such as future cost savings through reduced rates of medication errors and adverse drug events after the adoption of a CPOE system. As such, the extraction of the costs and outcomes accrued should extend beyond the implementation period to allow for the stabilization of the system to reach the point at which the change in outcomes is expected to occur.

The types of costs and outcomes to be included in an eHealth economic evaluation study should be clearly defined at the outset. The types of costs reported in the eHealth literature include one-time direct costs, ongoing direct costs, and ongoing indirect costs. Examples of one-time direct costs are hardware, software, conversion, training and support. Examples of ongoing direct costs

are system maintenance and upgrade, user/technical support and training. Examples of ongoing indirect costs are prorated IT management costs and changes in staff workload. The types of outcomes include revenues, cost savings, resource utilization, and clinical/health outcomes. Examples of revenues are money generated from billing and payment of services provided through the system and changes in financial arrangements such as reimbursement rates and accounts receivable days. Examples of labour, supply and capital savings are changes in staffing and supply costs and capital expenditures after system adoption. Examples of health outcomes are changes in patients' clinical conditions and adverse events detected. Note that the outcomes reported in the eHealth literature are mostly tangible in nature. There are also intangible outcomes such as patient suffering and staff morale affected by eHealth systems but they are difficult to quantify and are seldom addressed. For detailed lists of cost and outcome measures and references, refer to the additional online material (Appendices 9 and 10, respectively) in Bassi and Lau (2013).

14.3.2 Measurement of Costs and Outcomes

When measuring costs and outcomes, one needs to consider the costing approach, data sources and analytical methods used. Costing approach refers to the use of micro-costing versus macro-costing to determine the costs and outcomes in each eHealth system option (Roberts, 2006). Micro-costing is a detailed bottom-up accounting approach that measures every relevant resource used in system adoption. Macro-costing takes a top-down approach to provide gross estimates of resource use at an aggregate level without the detail. For instance, to measure the cost of a CPOE system with micro-costing, one would compile all of the relevant direct, indirect, one-time and ongoing costs that have accrued over the defined time period. With macro-costing, one may assign a portion of the overall IT operation budget based on some formula as the CPOE cost. While micro-costing is more precise in determining the detailed costs and outcomes for a system, it is a time-consuming and context-specific approach that is expensive and, hence, less generalizable than macro-costing.

The sources of cost and outcome data can be internal records, published reports and expert opinions. Internal records can be obtained retrospectively from historical data such as financial statements and patient charts, or prospectively from resource use data collected in a field study. Published reports are often publicly available statistics such as aggregate health expenditures reported at the regional or national level, and established disease prevalence rates at the community or population level. Expert opinions are ways to provide estimates through consensus when it is impractical to derive the actual detailed costs and outcomes, or to project future benefits not yet realized such as the extent of reduced medication errors expected from a CPOE system (Bassi & Lau, 2013, Table 4).

The analytical methods used to measure costs and outcomes can be based on accounting, statistical or operations research approaches. The accounting approach uses cost accounting, managerial accounting and financial accounting

methods to determine the costs and outcomes of the respective system options. The statistical approach uses such methods as logistic regression, general linear/mixed model and inferential testing for group differences (e.g., *t*-test, chi-square and odds ratio) to determine the presence and magnitude of the differences in costs and outcomes that exist among the options being considered. The operations research approach uses such methods as panel regression, parametric cost analysis, stochastic frontier analysis and simulation to estimate the direction and magnitude of projected changes in costs and outcomes for each of the options involved (Bassi & Lau, 2013, Table 4).

14.3.3 Valuation of Costs and Outcomes

Valuation is the determination of the monetary value of the costs and outcomes associated with the options being considered (Simoens, 2009). The key concepts in valuation when comparing the worth of each option are the notions of uncertainty, discounting, present value, inflation, and opportunity cost. These concepts are briefly outlined below.

- *Uncertainty* refers to the degree of imprecision in the costs and outcomes of the options. Such uncertainty can arise from the selected analytical methods, data samples, end point extrapolations and generalization of results. A common approach to handling uncertainty is through sensitivity analysis where a range of cost, outcome and other parameter estimates (e.g., time frame, discount rate) are applied to observe the direction and magnitude of change in the results (Brennan & Akehurst, 2000).

- *Discounting* is the incorporation of the time value of money into the costs and outcomes for each option being considered. It is based on the concept that a dollar is worth less tomorrow than today. Therefore discounting allows the calculation of the present value of costs and outcomes that can accrue differently over time. The most common discount rates found in the literature are between 3% and 5%. Often, a sensitivity analysis is performed by varying the discount rates to observe the change in results (Roberts, 2006).

- *Present value* (PV) is the current worth of a future sum of money based on a particular discount or interest rate. It is used to compare the expected cash flow for each of the options as they may accrue differently over time. A related term is net present value (NPV), which is the difference between the present value of the cash inflow and outflow in an option. When deciding among options, the PV or NPV with the highest value should be chosen (Roberts, 2006).

- *Inflation* is the sustained increase in the general price level of goods and services measured as an annual percentage increase called the inflation rate. In economic evaluation, the preferred approach is to use constant dollars and a small discount rate without inflation (known as the real discount rate). If the cost items inflate at different rates, the preferred approach is to apply different real discount rates to individual items without inflation (Drummond, Sculpher, et al., 2005).

- *Opportunity cost* is the foregone cost or benefit that could have been derived from the next best option instead of the one selected. When considering opportunity cost we are concerned with the incremental increases in healthcare budgets with alternative options and not the opportunity cost incurred elsewhere in the economy. One way to identify opportunity cost is to present healthcare and non-healthcare costs and benefits separately (Drummond, Sculpher, et al., 2005).

When attaching monetary values to costs and outcomes, one should apply current and locally relevant unit costs and benefits. The preference is to use published data sources from within the organization or region where the economic evaluation is done. If these sources are not available, then other data may be used but they should be adjusted for differences in price year and currency where appropriate. For discounting it should be applied to both costs and outcomes using the same discount rate. The reporting of undiscounted costs and outcomes should be included to allow comparison across contexts as local discount rates can vary. Where there is uncertainty in the costs and outcomes, sensitivity analysis should be included to assess their effects on the options (Brunetti et al., 2013).

14.3.4 Budget Impact and Priority Setting

Budget impact and priority setting relate to the overall importance of the respective investment decisions to the organization and its key stakeholder groups. In budget impact analysis, the focus is on the financial consequences of introducing a new intervention in a specific setting over a short to medium term. It takes on the perspective of the budget holder who has to pay for the intervention, with the alternative being the current practice, or status quo. In the analysis only direct costs are included typically over a time horizon of three years or less without discounting. For effectiveness, only short-term costs and savings are measured and the emphasis is on marginal return such as the incremental cost-effectiveness ratio that quantifies the cost for each additional unit of outcome produced. Sensitivity analysis is often included to demonstrate the impact of different scenarios and extreme cases (Garattini & van de Vooren, 2011).

In priority setting, program budgeting and marginal analysis is used to ensure optimal allocation of the limited resources available in the organization based on overall priorities. There are two parts to this analysis. The first part is program budgeting that is a compilation of the resources and expenditures allocated to existing services within the organization. The second part is marginal analysis where recommendations on investment of new services and disinvestment of existing services are made based on a set of predefined criteria by key stakeholders in the organization. An example is the multi-criterion decision analysis where a performance matrix is used to compare and rank options based on a set of policy-relevant criteria such as cost-effectiveness, disease severity, and affected population. The process should be supported by hard and soft evidence, and reflect the values and preferences of the stakeholder groups that are affected, for example the local population (Tsourapas & Frew, 2011; Baltussen & Niessen, 2006; Mitton & Donaldson, 2004).

14.4 Best Practice Guidance

The scoping review by Bassi and Lau (2013) of 42 published eHealth economic evaluation studies has found a lack of consistency in their design, analysis and reporting. Such variability can affect the ability of healthcare organizations in making evidence-informed eHealth investment decisions. At present there is no best practice guidance in eHealth economic evaluation, but there are two health economic evaluation standards that we can draw on for guidance. These are the Consensus on Health Economic Criteria (CHEC) list and the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist. They are described below.

14.4.1 CHEC List

The Consensus on Health Economic Criteria (CHEC) was published as a checklist to assess the methodological quality of economic evaluation studies in systematic reviews (Evers, Goossens, de Vet, van Tulder, & Ament, 2005). The list was created from an initial pool of items found in the literature, then reduced with three Delphi rounds by 23 international experts. The final list had 19 items, which are shown below (source: Table 1 in Evers et al., 2005, p. 243).

- Is the study population clearly described?
- Are competing alternatives clearly described?
- Is a well-defined research question posed in answerable form?
- Is the economic study design appropriate to the stated objective?

- Is the chosen time horizon appropriate to include relevant costs and consequences?
- Is the actual perspective chosen appropriate?
- Are all important and relevant costs for each alternative identified?
- Are all costs measured appropriately in physical units?
- Are costs valued appropriately?
- Are all important and relevant outcomes for each alternative identified?
- Are all outcomes measured appropriately?
- Are outcomes valued appropriately?
- Is an incremental analysis of costs and outcomes of alternatives performed?
- Are all future costs and outcomes discounted appropriately?
- Are all the important variables, whose values are uncertain, appropriately subjected to sensitivity analysis?
- Do the conclusions follow from the data reported?
- Does the study discuss the generalizability of the results to other settings and patient/client groups?
- Does the article indicate that there are no potential conflicts of interest of study researchers and funders?
- Are ethical and distributional issues discussed appropriately?

The authors emphasized that the CHEC list should be regarded as a minimal set of items when used to appraise an economic evaluation study in a systematic review. The additional guidance from the authors is: (a) having two or more reviewers and starting with a pilot when conducting the systematic review to increase rigour; (b) the items are subjective judgments of the quality of the study under review; and (c) journal publications should be accompanied by a detailed technical evaluation report.

14.4.2 CHEERS Checklist

The Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist was published in 2013 by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force (Husereau et al., 2013). Its purpose was to provide recommendations on the optimized reporting of health economic evaluation studies. Forty-four items were collated initially from the literature and reviewed by 47 individuals from academia, clinical practice, industry and government through two rounds of the Delphi process. A final list of 24 items with accompanying recommendations was compiled into six categories. They are summarized below.

- *Title and abstract* – two items on having a title that identifies the study as an economic evaluation, and a structured summary of objectives, perspective, setting, methods, results and conclusions.
- *Introduction* – one item on study context and objectives, including its policy and practice relevance.
- *Methods* – 14 items on target populations, setting, perspective, comparators, time horizon, discount rate, choice of health outcomes, measurement of effectiveness, measurement and valuation of preference-based outcomes, approaches for estimating resources and costs, currency and conversion, model choice, assumptions, and analytic methods.
- *Results* – four items on study parameters, incremental costs and outcomes, describing uncertainty in sampling and assumptions, and describing potential heterogeneity in study parameters (e.g., patient subgroups).
- *Discussion* – one item on findings, limitations, generalizability and current knowledge.
- *Others* – two items on source of study funding and conflicts of interest.

14.5 Exemplary Cases

This section contains three examples of eHealth economic evaluation studies that applied different approaches to determine the economic return on the investment made. The examples cover cost-benefit analysis, cost-effectiveness analysis, and simulation modelling. Readers interested in budget impact analysis may refer to the following:

- Fortney, Maciejewski, Tripathi, Deen, and Pyne (2012) on telemedicine-based collaborative care for depression.

- Anaya, Chan, Karmarkar, Asch, and Goetz (2012) on facility cost of HIV testing for newly identified HIV patients.

14.5.1 Cost-benefit of EMR in Primary Care

Wang and colleagues (2003) conducted a cost-benefit study to examine the financial impact of EMR on their organization in the ambulatory care setting. The *identified data sources* were cost and benefit data from the internal record, expert opinion and published literature. A five-year time horizon was used to cover all relevant costs and benefits. The *resource use measured* was the net financial cost or benefit per physician over five years. The *valuation of resource use* was the present value of net benefit or cost over five years based on historical data and expert estimates in 2002 U.S. dollars at a 5% discount rate.

The study findings showed the estimated net benefit was $86,400 per provider over five years. The benefits were from reduced drug expenditures and billing errors, improved radiology test utilization and increased charge capture. One-way sensitivity analysis showed net-benefit varied from $8,400 to $140,100 depending on the proportion of patients with care capitation. Five-way sensitivity analysis with most pessimistic and optimistic assumptions showed $2,300 net cost to $330,900 net benefit. This study showed EMR in primary care can lead to a positive financial return depending on the reimbursement mix.

14.5.2 Cost-effectiveness of Medication Ordering/Administration in Reducing Adverse Drug Events

Wu, Laporte, and Ungar (2007) conducted a cost-effectiveness study to examine the costs of adopting a medication ordering and administration system and its potential impact on reducing adverse drug events (ADEs) within the organization. The *identified data sources* were system and workload costs from internal records and expert opinion, and estimated ADE events from the literature. The *resource use measured* were annual cost and ADE rate projected over 10 years. The *valuation of resource use* was the annual system and workload costs based on historical data and expert estimates as net present value in 2004 Canadian and U.S. dollars at 5% discount rates.

The study findings showed the incremental cost-effectiveness of the new system was $12,700 USD per ADE prevented. Sensitivity analysis showed cost-effectiveness to be sensitive to the ADE rate, cost of the system, effectiveness of the system, and possible costs from increased physician workload.

14.5.3 Simulation Modelling of CPOE Implementation and Financial Impact

Ohsfeldt et al. (2005) conducted a simulation study on the cost of implementing CPOE in rural state hospitals and the financial implications of statewide implementation. The *identified data sources* included existing clinical information

system (CIS) status from a hospital mail survey, patient care revenue and hospital operating cost data from the statewide hospital association, and vendor CPOE cost estimates. The *resource use measured* was the net financial cost or benefit per physician over five years. The *valuation of resource use* was the operating margin present value of net benefit or cost over five and 10 years based on historical data and expert estimates in 2002 U.S. dollars at a 5% discount rate. Quadratic interpolation models were used to derive low and high cost estimates based on bed size and CIS category. Comparison of operating margins for first and second year post-CPOE across hospital types was done with different interest rates, depreciation schedules, third party reimbursements and fixed/marginal cost scenarios.

The study findings showed CPOE led to substantial operating costs for rural and critical access hospitals without substantial cost savings from improved efficiency or patient safety. The cost impact was less but still dramatic for urban and rural referral hospitals. For larger hospitals, modest benefits in cost savings or revenue enhancement were sufficient to offset CPOE costs. In conclusion, statewide CPOE adoption may not be financially feasible for small hospitals without increased payments or subsidies from third parties.

14.6 Implications

The eHealth economic evaluation methods described in this chapter have important implications for policy-makers and researchers involved with the planning, adoption and evaluation of eHealth systems. First, it is important to have a basic understanding of the principles and application of different eHealth economic evaluation methods as their selection is often based on a variety of contexts, perspectives and assumptions. Second, when conducting an eHealth economic evaluation it is important to be explicit in describing the identification, measurement and valuation steps to ensure all of the important and relevant costs and outcomes are included and handled appropriately. Third, to ensure rigour and to increase the generalizability of the eHealth economic evaluation study findings, one should adhere to the best practice guidance in their design, analysis and reporting.

To ensure rigour one should be aware of and avoid the common "methodological flaws" in the design, analysis and reporting of economic evaluation studies, as cautioned by Drummond and Sculpher (2005). The common design flaws are the omission of important and relevant costs and outcomes and the inclusion of inappropriate options for comparison, such as unusual local practice patterns in usual care, which can lead to incomplete and erroneous results. The common flaws in data collection and analysis are the problems of making indirect clinical comparisons, inadequate representation of the underlying effectiveness data, inappropriate extrapolation beyond the time period of the study, over-reliance on assumptions, and inadequate handling of uncertainty. For instance, the presence of major baseline group differences across the options

would make the results incomparable. The common flaws in reporting are the inappropriate aggregation of results, inclusion of only the average cost-effectiveness ratios, inadequate handling of generalizability, and selective reporting of the findings. In particular, the reporting of average cost-effectiveness ratios based on total costs divided by total effects is common in the eHealth literature and can be misleading since it does not show the incremental cost involved to produce an extra unit of outcome.

The generalizability of eHealth economic evaluation study findings can be increased by drawing on the recommendations of the National Health Service Health Technology Assessment Programme in the United Kingdom on the design, analysis and reporting of economic evaluations (Drummond, Manca, & Sculpher, 2005). For trial-based studies, the design should ensure the representativeness of the study sites and patients, the relevance of the options for comparison, the ability to include different perspectives, the separation of resource use data from unit costs or pricing, and the use of health state preferences that are relevant to the populations being studied. The analysis of multi-location/centre trials should test for the homogeneity of the data prior to pooling of the results to avoid the clustering of treatment effects. The reporting of trial-based results should include the characteristics of the study sites supplemented with a detailed technical report to help the readers better understand the contexts and decide if the findings are relevant to their organizations.

For model-based studies, the design should be clear in specifying the decision problem and options, identifying the stakeholders to be informed by the decision model, and ensuring the modelling approaches are relevant to the stakeholders (e.g., the perspective and objective function). The analysis of model-based trials should justify its handling of the cost, resource use, effectiveness and preference value data, especially when there is uncertainty and heterogeneity in the data across groups, locations and practices. The reporting of model-based results should include the justifications of the parameter inputs to the model to ensure they are appropriate and relevant to the stakeholders. Any pre-analysis done on the input data so they can be incorporated into the model should be explained to justify its relevance.

14.7 Summary

This chapter described the different methods that are used in eHealth economic evaluation. The methods cover different analytical approaches and the process for resource costing and determining the outcomes. There are also published best practice standards and guidelines that should be considered in the design, analysis and reporting of eHealth economic evaluation studies. The three case studies provide examples of how the economic evaluation of eHealth systems is done using select methods.

References

Anaya, H. D., Chan, K., Karmarkar, U., Asch, S. M., & Goetz, M. B. (2012). Budget impact analysis of HIV testing in the VA healthcare system. *Value in Health, 15*(8), 1022–1028.

Baltussen, R., & Niessen, L. (2006). Priority setting of health interventions: the need for multi-criteria decision analysis. *Cost Effectiveness and Resource Allocation, 4*, 14. doi: 10.1186/1478-7547-4-14

Bassi, J., & Lau, F. (2013). Measuring value for money: A scoping review on economic evaluation of health information systems. *Journal of American Medical Informatics Association, 20*(4), 792–801.

Brennan, A., & Akehurst, R. (2000). Modelling in health economic evaluation. What is its place? What is its value? *Pharmacoeconomics, 17*(5), 445–459.

Brunetti, M., Shemilt, I., Pregno, S., Vale, L., Oxman, A. D., Lord, J., … Schunemann, H. J. (2013). GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. *Journal of Clinical Epidemiology, 66*(2), 140–150.

Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddart, G. L. (2005). *Methods for the economic evaluation of health care programmes* (3rd ed.). Oxford: Oxford University Press.

Drummond, M., & Sculpher, M. (2005). Common methodological flaws in economic evaluations. *Medical Care, 43*(7 suppl), 5–14.

Drummond, M., Manca, A., & Sculpher, M. (2005). Increasing the generalizability of economic evaluations: Recommendations for the design, analysis and reporting of studies. *International Journal of Technology Assessment in Health Care, 21*(2), 165–171.

Evers, S., Goossens, M., de Vet, H., van Tulder, M., & Ament, A. (2005). Criteria list for assessment of methodological quality of economic evaluations: Consensus on health economic criteria. *International Journal of Technology Assessment in Health Care, 21*(2), 240–245.

Fortney, J. C., Maciejewski, M. L., Tripathi, S. P., Deen, T. L., & Pyne, J. M. (2011). A budget impact analysis of telemedicine-based collaborative care for depression. *Medical Care, 49*(9), 872–880.

Garattini, L., & van de Vooren, K. (2011). Budget impact analysis in economic evaluation: a proposal for a clearer definition. *European Journal of Health Economics, 12*(6), 499–502.

Husereau, D., Drummond, M., Petrou, S., Carswell, C., Moher, D., Greenberg, D., … Loder, E. (2013). Consolidated health economic evaluation reporting standards (CHEERS) – Explanation and elaboration: A report of the ISPOR health economic evaluation publication guidelines good reporting practices task force. *Value in Health, 16*, 231–250.

Mitton, C., & Donaldson, C. (2004). Health care priority setting: principles, practice and challenges. *Cost Effectiveness and Resource Allocation, 2*, 3. doi: 10.1186/1478-7547-2-3

Ohsfeldt, R. L., Ward, M. M., Schneider, J. E., Jaana, M., Miller, T. R., Lee, Y., & Wakefield, D. S. (2005). Implementation of hospital computerized physician order entry systems in a rural state: Feasibility and financial impact. *Journal of American Medical Informatics Association, 12*(1), 20–27.

Roberts, M. S. (2006). Economic aspects of evaluation. In C. P. Friedman & J. C. Wyatt (Eds.), *Evaluation methods in biomedical informatics* (2nd ed., pp. 301–337). New York: Springer.

Simoens, S. (2009). Health economic assessment: a methodological primer. *International Journal of Environmental Research and Public Health, 6*(12), 2950–2966.

Tsourapas, A., & Frew, E. (2011). Evaluating "success" in programme budgeting and marginal analysis: a literature review. *Journal of Health Services Research & Policy, 16*(3), 177–183.

Wang, S. J., Middleton, B., Prosser, L. A., Bardon, C. G., Spurr, C. D., Carchidi, P. J., … Bates, D. W. (2003). A cost-benefit analysis of EMR in primary care. *American Journal of Medicine, 114*(5), 387–403.

Wu, R. C., Laporte, A., & Ungar, W. J. (2007). Cost-effectiveness of an electronic medication ordering and administration system in reducing adverse drug events. *Journal of Evaluation in Clinical Practice, 13*(3), 440–448.

Appendix

Glossary of Terms

Economic Analysis	Description (based on Roberts, 2006; Chisholm, 1998; Robinson, 1993).
Cost-minimization Analysis	Costs are measured in dollars and outcomes are assumed to be equivalent. Purpose of this analysis is to determine the least cost option.
Cost-consequence Analysis	Costs are measured in dollars and outcomes are measured in variable and multiple units. This analysis lists the individual outcomes without further aggregation.
Cost-effectiveness Analysis	Costs are measured in dollars and outcomes are measured in clinical terms or natural units. This analysis uses a common unit of outcome to express the cost of each option.
Cost-utility Analysis	Costs are measured in dollars and outcomes are measured as utility (subjective satisfaction). A common utility measure is quality-adjusted life year (QALY).
Cost-benefit Analysis	Costs are measured in dollars and outcomes are measured in dollars. This analysis is used to assess which option is best based on monetary values for costs and benefits. In generally, benefits should exceed costs for an option to be worthwhile.

Common Analytical Measures		
Analytical Term	Description	Sources
ANOVA	A statistical procedure to test if differences exist among two or more groups of subjects on one or more factors.	Dawson and Trapp (2004, p. 403)
Average cost	Cost of producing one unit of output.	Drummond et al. (2005, p. 65)
Chi-square	A statistical procedure to test if the proportions of two or more factors are equal which suggests they are independent of each other.	Dawson and Trapp (2004, p. 404)
Cost amortization/ depreciation	Spreading the cost of an intangible/tangible asset over a fixed period that represents the useful life of that asset.	Haber (2008, p. 86)
Cost savings	Action that will result in fulfillment of the objectives of a purchase at a cost lower than the historical cost or the projected cost.	Online Business Dictionary (n.d.)
Discounting	Process of finding the present value of an amount or series of cash flows expected in the future.	Gapenski (2009, p. 255)
Discount rate	The real rate of return, or interest rate, that will be returned in the future on the money invested today rather than being spent.	Roberts (2006, p. 320)
Incremental cost-benefit ratio (ICBR)	A ratio of the net cost of implementing one system over another divided by the net benefit, measured in monetary term. The unit is expressed as the cost of an additional unit of money generated as the benefit.	Simoens (2009, p. 2596)
Incremental cost-effectiveness ratio (ICER)	A ratio of the net cost of implementing one system over another divided by the net benefit, measured as a clinical outcome. The unit is expressed as the cost of an additional unit of a given outcome measure as the benefit.	Roberts (2006, p. 316)

Common Analytical Measures		
Incremental cost-utility ratio (ICUR)	A ratio of the net cost of implementing one system over another divided by the net benefit, measured as a health utility such as quality-adjusted life years. The unit is expressed as the cost of an additional unit of a given health utility measure as the benefit.	Simoens (2009, p. 2596)
Inflation	Change in prices over time within an economy that needs to be standardized to a common base year if the costs span multiple years.	Roberts (2006, p. 320)
Least cost	The option where the cost is minimized with the most quantity of outcome.	Roberts (2006, p. 315)
General linear model	A statistical model used to predict the outcome from a set of independent variables.	Dawson and Trapp (2004, p. 192)
Generalized linear mixed model (GLMM)	A form of regression analysis of correlated data from subjects with multiple longitudinal responses in the data set based on a logit link function.	Cnaan et al. (1997)
Logistic regression	A technique to predict an outcome from one or more independent variables when the outcome is a binary variable.	Dawson and Trapp (2004, p. 408)
Markov chain	A simulation modelling technique to determine the probability of an event going from one state to the next.	Ravindran (2008, chapter 8)
Mean inefficiency score	The per cent difference between the cost of an organization and the frontier determined by the aggregate cost of all organizations using stochastic frontier analysis.	Carey et al. (2008)
Monte Carlo simulation	Statistical modelling techniques that emulate the behaviour and performance of a system as events take place over time.	Ravindran (2008, chapter 8)
Net benefit	Also known as net monetary benefit, which is the difference between the amount an organization is willing to pay for the increase in effectiveness and the increase in cost.	Drummond et al. (2005, p. 131)
Net present value (NPV)	The dollar value of an investment discounted at the opportunity cost of capital.	Gapenski (2009, p. 258)
Operating margin	Amount of operating profit per dollar of operating revenues. Also referred to as the proportion of revenue left over after paying for variable costs of production in order to pay for fixed costs such as interests on debt.	Gapenski (2009, p. 381)
Panel regression, fixed effect	A regression technique that uses two-dimensional panel data collected over time on the same subjects that have unique attributes not due to random variations.	Baltagi (2011)
Parametric cost analysis	A cost estimating technique that uses regression methods to develop cost estimating relationships to establish cost estimates with one or more independent variables.	AcqNotes (n.d.)
Payback	Number of years that it takes to recover the cost of an investment.	Gapenski (2009, p. 248)
Quality-adjusted life year (QALY)	The period of time in perfect health that a patient says is equivalent to a year in a state of ill health.	Sox et al. (2006, p. 217)
Regression	A technique to predict an outcome from one or more independent variables.	Dawson and Trapp (2004, p. 412)
Regression coefficient	The slope of the regression line in a simple linear regression, or the weights applied to independent variables in multiple regression.	Dawson and Trapp (2004, p. 412)
Return on investment (ROI)	Profitability of an investment, measured in dollars or rate of return.	Gapenski (2009 p. 258)

Common Analytical Measures		
Sensitivity analysis	A technique to test the stability of the outputs of an analysis over a range of input variable estimates.	Sox et al. (2006, p. 163)
Scenarios analysis	A series of alternative cases with variable estimates that represent the realistic, best and worst cases to be considered in the analysis.	Drummond et al. (2005, p. 43)
Stochastic frontier analysis	An economic modelling technique that estimates production or cost functions while taking into account the inefficiency that exists within the organization.	Online Encyclopaedia (n.d.)
t-test	A statistical test to compare a mean with a norm or two means with small sample sizes.	Dawson and Trapp (2004, p. 414)

References for Appendix

AcqNotes. (n.d.). Retrieved from http://www.acqnotes.com/Tasks/Parametric%20Cost%20Estimating%20.html

Baltagi, B. H. (2011). *Econometrics* (5th ed.). New York: Springer.

Carey, K., Burgess, J. F., & Young, G. J. (2008). Specialty and full service hospitals: A comparative cost analysis. *Health Research and Educational Trust, 43*(5, Part II), 1869–1887. doi: 10.1111/j.1475-6773.2008.00881.x

Chisholm, D. (1998). Economic analyses. *International Review of Psychiatry, 10*(4), 323–330.

Cnaan, A., Laird, N. M., & Slasor, P. (1997). Using the general linear mixed model to analyse unbalanced repeated measures and longitudinal data. *Statistics in Medicine*, *16*(20), 2349–2380.

Dawson, B., & Trapp, R. G. (2004). *Basic and clinical biostatistics* (4th ed.). New York: Lange Medical Books/McGraw-Hill.

Drummond, M. F., Sculpher, M. J., Torrance, G. W., O'Brien, B. J., & Stoddart, G. L. (2005). *Methods for the economic evaluation of health care programmes* (3rd ed.). Oxford: Oxford University Press.

Gapenski, L. C. (2009). *Fundamentals of healthcare finance*. Chicago: Health Administration Press.

Haber, J. R. (2008). *Accounting demystified*. New York: American Management Association/Amacom.

Online Business Dictionary. (n.d.). Fairfax, VA: WebFinance Inc. Retrieved from http://www.BusinessDictionary.com

Online Encyclopaedia. (n.d.). Boston: Cengage Learning. Retrieved from http://www.encyclopedia.com

Ravindran, R. A. (Ed). (2008). *Operations research and management science handbook*. NewYork: CRC Press/Taylor & Francis Group.

Roberts, M. S. (2006). Economic aspects of evaluation. In C. P. Friedman & J. C. Wyatt (Eds.), *Evaluation methods in biomedical informatics* (2nd ed., pp. 301–337). New York: Springer.

Robinson, R. (1993). Economic evaluation and health care: What does it mean? *British Medical Journal, 307*(6905), 670–673.

Simoens, S. (2009). Health economic assessment: a methodological primer. *International Journal of Environmental Research and Public Health, 6*(12), 2950–2966.

Sox, H., Blatt, M. A., Higgins, M. C., & Marton, K. T. (2006). *Medical decision making* (1st ed.). Boston: Butterworth-Heinemann.

Chapter 15

Methods for Modelling and Simulation Studies

James G. Anderson, Rong Fu

15.1 Introduction

Evaluation of the implementation and use of an eHealth System such as electronic health records (EHR), decision support systems, computerized provider order entry and telehealth frequently require the use of methods other than traditional randomized control trials. Moehr (2002) points out some of the problems involved in evaluating eHealth applications. He suggests that these evaluations need to include the dynamic process of adaptation of the system and its environment rather than just its technical features. Conventional evaluation methods do not adequately describe the dynamic nature of eHealth systems.

Regression analysis, network analysis and computer simulation provide alternative methodologies that help investigators better understand the impact of these systems on workflow, cost, effectiveness, and quality of healthcare delivery. The analytical approaches described below focus on different aspects of eHealth systems. Regression analysis examines attribute data to answer such questions as which physician characteristics predict EHR use. Network analysis explores relationships among members of a network, such as a medical practice, to determine how communication among members affects use of clinical practice guidelines. The focus of simulation is on the system level to explore issues such as how to alleviate crowding in the Emergency Department (ED). In general, the chapter is aimed at practitioners with little or no experience in designing and implementing these methods.

15.2 Regression Analysis

Regression analysis is a statistical tool that attempts to estimate an outcome (also known as the response variable) based on a set of predictors (also known as the explanatory variables). Specifically, a regression model explores how the typical value of the response variable changes given different values of the explanatory variable(s). There are several regression methods used widely in quantitative research: linear, logistic, and multivariate regression models. Apart from these common regression models, time series regression and structural equation modelling are relatively new regression tools in eHealth studies.

Regression models allow explanations and predictions of past, present, or future events with information obtained from internal or external sources. Regression analysis can be performed with both cross-sectional data and panel data. Cross-sectional data are collected by observing many subjects (e.g., individuals, hospitals) at a particular point in time. Panel data, also called longitudinal data or cross-sectional time series data, are collected by observing the same subjects at two or more time periods. In order to build a regression model, one needs to determine the response variable(s), the explanatory variable(s), the time frame, and the specific analytical model.

15.2.1 Types of Regression Models

Linear regression is the most basic and commonly used technique for determining how the response variable is affected by changes in one or more explanatory variables. Whereas a simple linear regression model predicts the outcome based on a single explanatory variable, a multiple linear regression model uses two or more explanatory variables to predict the response variable. In linear regression analysis, the relationship between the predictor(s) and the outcome is typically plotted as a straight line that best approximates all the individual data points. A possible research question that can be answered using linear regression is the following: What is the association between eHealth literacy (the ability to seek and understand health information from electronic sources) and colorectal cancer knowledge (see Mitsutake, Shibata, Ishii, & Oka, 2012)?

Logistic regression is an extension of linear regression that allows one to predict categorical outcomes based on predictors. A categorical outcome is one that takes on one of a fixed number of possible values (e.g., the blood type of a person has four categories: A, B, AB or O). In eHealth evaluation, a logistic regression model is commonly used to model the linear relationship between a binary outcome variable (a categorical variable with only two values) and one or more predictors. The binary outcome variable usually takes the value of 0 or 1 to indicate the absence or presence of an outcome (e.g., 0 = survival, 1 = death). Thereby, logistic regression models are widely used to predict the odds of the presence of the outcome based on the values of the predictors. A possible research question that can be answered using logistic regression is the following: Do eHealth literacy and patient-centred communication affect the odds of post-visit online health information seeking (see Li, Orrange, Kravitz, & Bell, 2014)?

A multivariate regression model estimates more than one outcome based on a set of predictors. This model attempts to determine a formula that describes how elements in a set of variables respond simultaneously to changes in others. The main characteristic that distinguishes multivariate regression from multiple regression is the use of multiple outcomes. A possible research question that can be answered using multivariate regression is the following: What is the relationship between basic electronic medical records and outcomes such as having a patient safety event, impatient death, and hospital readmission (see Encinosa & Bae, 2011)?

A time series regression model predicts a future outcome based on the outcome history and the transfer of dynamics from a series of predictors. In order to use this model, one needs to have measurements that are taken from the same subjects at successive time points (e.g., hospital readmission rates in five separate years). A possible research question that can be answered using time series regression analysis on panel data is the following: How do hospital information technologies affect hospital operating expenses across three years (see Bardhan & Thouin, 2013)?

Structural equation modelling (SEM) is a family of related statistical procedures designed to determine and validate a proposed process and/or a theoretical model. SEM can be used to examine research questions involving the indirect or direct observation of one or more predictors and/or one or more outcomes. Some common SEM methods include confirmatory factor analysis, path analysis, and latent growth modelling (Kline, 2010).

Confirmatory factor analysis is a multivariate statistical procedure used to verify the hypothesized relationship between observed variables and their underlying latent constructs. The eHealth literacy study presented by Neter and Brainin (2012) is a good example of confirmatory factor analysis. Path analysis is an extension of multiple regression that evaluates causal models by examining the relationship between one or more explanatory and response variables. A case in point is that Cho, Park, and Lee (2014) used a path analysis to examine the effects of several cognitive factors (e.g., health consciousness, health information orientation) on the extent of health-app use. Latent growth modelling is a longitudinal analysis technique that can estimate growth over a period of time. Anderson, Ramanujam, Hensel, and Sirio (2010) used latent growth curve analysis to examine longitudinal trends in the quarterly number of errors and associated corrective actions reported by 25 hospitals.

15.2.2 Evaluating Electronic Medical Records using Multivariate Regression

Encinosa and Bae (2011) used multivariate regression models to examine whether electronic medical records (EMRs) contain costs in the Patient Protection (PP) and Affordable Care Act (ACA) reforms to reduce patient safety events. In this study, data were obtained from the 2007 MarketScan Commercial Claims and Encounter Database, the 2007 American Hospital

Association Annual Survey and its Information Technology Supplement. The methodological components for this study are summarized as follow:

- *Research question #1* – What is the relationship between basic EMRs and the probability that a surgery will have a patient safety event?
  - *Outcome #1* – Patient safety event, measured by "surgical-related patient safety events" with 12 indicators, "nursing-related patient safety events" with 5 indicators, and other "likely preventable patient safety events" with 7 indicators.

- *Research question #2* – What is the relationship between basic EMRs and the probability of inpatient death within 90 days following surgery?
  - *Outcome #2* – Death, measured by any inpatient hospital death occurring within 90 days following surgery.

- *Research question #3* – What is the relationship between basic EMRs and the probability of a 90-day readmission for surgeries?
  - *Outcome #3* – Readmission, measured by any overnight stays at an inpatient hospital within 90 days following surgery.

- *Research question #4* – What is the relationship between basic EMRs and total 90-day hospital expenditures?
  - *Outcome #4* – Hospital expenditures, measured by transacted prices including all inpatient hospital, physician, drug, and lab payments for any inpatient stay occurring up to 90 days following surgery.

- *Analytical model* – Multivariate regression models.

- *Time frame* – A cross-sectional design where the data were collected all at the same time or within a short time frame.

- *Predictors* – Basic EMRs, a binary variable (1 = having basic EMRs; 0 = no basic EMRs) measured by whether a hospital has the following eight basic EMR functionalities in at least one major clinical unit: demographic characteristics of patients, problem lists, medication lists, discharge summaries, laboratory reports, radiologic reports, diagnostic test results, and computerized provider order entry for medications.

- *Covariates* – Age, sex, suffering from hypertension, suffering from diabetes, suffering from liver disease, suffering from depression, obesity, etc.

The study findings showed that EMRs did not reduce the rate of patent safety events. However, once a patient safety event occurs, EMRs reduced death by 34%, readmissions by 39%, and hospital expenditures by $4,840 (16%). These results were obtained by examining the relationships between multiple outcomes and predictors in multivariate regression models after controlling for covariates. Taken together, the findings of this study indicate that EMRs contain costs in the PP and ACA reforms by better coordinating care to rescue patients from medical errors once a patient safety event occurs.

15.2.3 Evaluating Health Information Technologies using Time Series Regression on Panel Data

Bardhan and Thouin (2013) applied time series regression models to panel data to estimate the impact of health information technologies (HIT) on hospital operating expenses and the quality of healthcare delivery during the three-year period. In this study, data on hospital information technologies usage was obtained from the Dorenfest Institute for Health Information Technology Research database. Data on hospital process quality measures was obtained from the U.S. Department of Health and Human Services (HHS) Hospital Compare Program. Data on hospital operating expenses was obtained using publicly available data from the U.S. Center for Medicare and Medicaid Services. The methodological components of this study are summarized as follow:

- *Research question #1* – What is the relationship between implementation of HIT and the quality of healthcare delivery indicated by levels of conformance to evidence-based best practices?
  - *Outcome #1* – Acute myocardial infarction, with eight process quality measures.
  - *Outcome #2* – Heart failure, with four process quality measures.
  - *Outcome #3* – Pneumonia, with seven process quality measures.
  - *Outcome #4* – Surgical infection prevention, with two process quality measures.

- *Research question #2* – What is the relationship between implementation of HIT and hospital operating expenses?
  - *Outcome #5* – Operating expense per bed, measured by dividing the hospitals' operating costs for providing healthcare services by the total number of beds in use.

- *Analytical model* – Time series regression on panel data.

- *Time frame* – A three-year longitudinal design where data were collected each year from 2004 to 2006.

- *Predictors* – Clinical systems (six factors), financial systems (four factors), scheduling systems (one factor), and human resources systems (two factors).

- *Covariates* – Hospital type, hospital size, hospital case mix index, hospital location, and teaching status.

The study findings indicated that usage of clinical information systems and patient scheduling applications was associated with greater conformance with best practices for treatment of heart attacks, heart failures, and pneumonia. Whereas financial and human resource management systems were associated with lower hospital operating expenses, implementation of clinical information systems and scheduling systems was associated with higher operating expenses. Taken together, the findings of this study suggest that investments in HIT have a positive impact on the overall quality of healthcare delivery. However, the effect of HIT implementation on hospital operating expenses is mixed and needs to be factored into consideration when making implementation decisions.

15.3 Social Network Analysis

Social network analysis comprises a set of methods that can be used to investigate patterns of relationships among individuals, departments, organizations, etc. These relationships affect behaviour such as adoption and use of electronic medical records, decision support systems, and telehealth (Anderson, 2002a).

15.3.1 Social Networks and Physician Adoption of Electronic Health Records

Zheng, Padman, Krackhardt, Johnson, and Diamond (2010) studied how social interactions influence physician adoption of EHRs. A survey was used to identify social interactions among 40 residents and 15 attending physicians in an ambulatory care primary care practice. Social network analysis was used to determine the relation of the structure of interactions to physicians' rates of utilization of the EHR.

- *Objective* – To examine how social influences affect physician EHR adoption.

- *Research Hypothesis #1* – The level of EHR adoption can be predicted by cohesion over the professional network, the friendship network and the perceived influence network among physicians. Cohesion reflects how well physicians were connected to each

other and whether key individuals possess pivotal positions in the network.

- *Research Hypothesis #2* – The level of EHR adoption can be predicted by structural equivalence of the professional network, the friendship network and the perceived influence network. Structural equivalence measures the similarity in interaction patterns in the three types of networks.

- *Data Collection* – A social network survey was administered to 55 physicians affiliated with an outpatient primary care practice associated with a 512-bed tertiary care medical facility. The survey asked physicians: (a) to name their colleagues that they consulted with on patient care issues; (b) which colleagues they considered to be personal friends; and (c) which colleagues influenced them to use the EHR. A second survey assessed personal characteristics such as gender, work experience, computer literacy, attitude toward use of the EHR, etc.

- *Outcome Measures* – Rates of EHR usage for patient data documentation or retrieval of patient data were calculated for each physician.

- *Analysis* – The analysis assessed the influence of the social structure and structural equivalence on rates of EHR system usage.

Results of the analysis indicated that several physicians provided the bulk of information concerning patient care in the professional network. In contrast, analysis of the perceived influence network suggested that influence over adoption of the EHR rarely occurred in the clinic. Analysis of the friendship network indicated that residents who had named the same attending physician as a personal friend exhibited comparable EHR adoption behaviour.

The results of this study suggest that identifying opinion leaders who developed friendships with many other members of a medical practice can be used to promote the diffusion of innovations like EHRs.

15.3.2 The Use of Social Networks to Study Outbreaks of Hospital-acquired Infections

Cusumano-Towner, Li, Tuo, Krishnan, and Maslove (2013) used social network analysis to study outbreaks of nosocomial infections among hospital patients. EMR data were used to model contacts among patients through shared rooms and contact with healthcare workers. The social networks were used to conduct probabilistic simulations of outbreaks of Methicillin-resistant Staphylococcus aureus (MRSA) and influenza.

- *Objective* – The objectives of this study were: (a) to create a social network of hospital patients using data from an EMR; (b) to use the network to simulate nosocomial outbreaks of MRSA and influenza; and (c) to identify potential interventions.

- *Data* – EMR data were extracted from a clinical data warehouse covering hospital admissions over a 70-day period. Data from days 35 to 45 were used in the simulation. Shared contact with healthcare workers was determined from metadata contained in clinical documents.

- *Analysis* – The data files were used to construct networks of pairwise connections between individual patients based on sharing of rooms and shared contact with healthcare workers. The two networks were combined into a graph of epidemiologic links that change over time. This social network was used to develop a probabilistic model of the spread of infection through the hospital. The probabilistic model was used to simulate outbreaks of MRSA and influenza and to test the potential effects of infection control measures. Infections originating in the ED, a medical step-down unit, and a psychiatry unit were simulated.

The results indicated that the risk of spreading influenza between wards was greatest between two psychiatric units, and between the cardiac unit and coronary care unit. The ED and operating areas had low levels of incoming infection. Its simulations predicted that vaccination of the staff could markedly decrease the spread of influenza. Simulation of outbreaks of MRSA predicted that an infection originating in the medical step-down unit spread to the ICU, the neurosurgical, orthopedic, and cardiac units. The risk of transmission of MRSA was substantially mitigated by a 50% increase in hand hygiene compliance. The benefits of the approach used in this study are: First, it used existing data collected during clinical care and stored in an EMR to construct patient networks; second, these data reflect local staffing and patient flow patterns unique to the hospital under investigation; third, this approach allows for real-time updating of the patient networks; and fourth, social networks can be used to model the effects of infection control interventions such as patient isolation, hand hygiene, and staff vaccination.

15.4 Simulation Modelling

The development of a computer simulation model begins with a system analysis. Important elements of the system and relationships among them are identified. Data used in defining the system may be obtained from system logs, interviews, questionnaires, work sampling and expert judgment (Anderson, 2002b, 2002c).

There are several types of simulation: discrete event, continuous, and agent-based. In a discrete event model, items (e.g., patients, medical orders, etc.) flow through a network of components. Each component performs a function (e.g., MRI) before the item (e.g., patient) moves on to the next component (e.g., service). For a discrete event simulation of a computerized physician order entry system, see Anderson et al. (1988).

Continuous simulation is used when an eHealth system involves a continuous flow of information, patients, material, or other resources. The model is comprised of state variables (e.g., the number of patients in the system at any time), rates of flow (e.g., entry of new patients and exit of existing patients), and control variables that affect the flow rates. For a model of a drug ordering and delivery system of a hospital, see Anderson, Jay, Anderson, and Hunt (2002). Continuous simulation models such as systems dynamics are comprised of a set of differential equations representing feedback loops among state variables that represent the system under investigation. This feedback structure is what makes the system adapt over time.

Agent-based models are used to determine the global consequences of interactions among individual agents. Agents generate emergent behaviour by interacting with one another according to a small set of rules. Interactions among agents give rise to the system's behaviour. For an agent-based model of the healthcare system of a refugee community, see Anderson, Chaturvedi, and Cibulskis (2007).

Once a simulation model has been constructed, it is validated against historical data that describes the behaviour of the system over time. A major advantage of simulation is that the model can be used to make modifications (e.g., the number of RNs or MDs in the ER) and predict effects on the system's performance. Such computer experiments can be performed without disrupting the practice setting.

15.4.1 Forecasting Emergency Department Crowding using Discrete Event Simulation

Hoot et al. (2009) applied discrete event simulation to forecasting emergency department crowding. The growing problem of crowding in emergency departments is resulting in delayed treatment, prolonged transport, increased mortality, and financial burdens on hospitals. This study developed and validated a method of forecasting future emergency department crowding using discrete event simulation.

- *Objective* – Implement and validate a simulation model to be used in forecasting future crowding in emergency departments.

- *Research question #1* – Could a simulation model accurately predict future crowding based on existing data from emergency department information systems?

- *Research question #2* – How well does the model predict future values of several crowding measures in a real operational setting?

- *Methods* – A discrete event simulation model was constructed and validated based on data from an adult ED in a tertiary care, urban Level 1 trauma center. The model describes patient arrivals, evaluation, treatment and potential hospital admissions.

- *Input variables* – The following data were collected in an adult emergency department of a tertiary care medical centre during a three-month period:
  - Time of initial registration in the ED.
  - Time placed in an ED bed.
  - Time of request for a hospital bed.
  - Time of discharge from the ED.
  - Patient's triage category.
  - Whether the patient left the ED without being seen.

- *Outcome measures* – The model forecasts the following crowding measures:
  - Number of patients in the waiting room.
  - Average waiting time.
  - Occupancy – total number of patients in ED beds.
  - Length of stay in the ED.
  - Number of patients awaiting hospital admission.
  - Average time patients waited for hospital admission.
  - Probability of ambulance diversion due to ED crowding.

The simulation model provides accurate real-time forecasts of inputs, throughputs and output measures of crowding up to eight hours in the future. The tool could be used in other EDs that have information systems that provide the six patient-level variables.

15.4.2 Preventing Adverse Drug Events using Continuous Simulation

Anderson, Jay, Anderson, and Hunt (2002) developed a computer simulation model to evaluate information technology applications designed to detect and prevent hospital medication errors that may result in adverse drug events. Model parameters were estimated from a study of prescription errors on two hospital medical/surgical units and used for the baseline simulation. The study evaluated five prevention strategies.

- *Objective* – To develop a model that can be used to evaluate the effectiveness of IT applications designed to prevent medical errors that may result in adverse drug events (ADEs) in hospitals.

- *Research Question #1* – How effective are each of five interventions in reducing ADEs in a hospital?

- *Research Question #2* – How effective are each of the interventions in reducing additional days of hospitalization that result from ADEs?

- *Research Question #3* – How effective are each of the five interventions in reducing the cost resulting from ADEs?

- *Methods* – A computer simulation model was constructed to represent the medication delivery system in a hospital. STELLA, a continuous simulation software package, was used to construct the model. Parameters of the model were estimated from a study of prescription errors on two hospital medical/surgical units.

- *Input Variables* – The following variables were obtained from a study of two hospital units:
  - Number of medication orders entered into the hospital information system.
  - Number and type of errors made in writing prescriptions.
  - Severity of medication errors.
  - Rates of ADEs resulting from medication errors.
  - Rates of errors committed during the dispensing and administration of medications were based on published studies.

- *Interventions* – The model was used to evaluate the following interventions:
  - Provision of drug information by the Hospital Information System when prescriptions are written.
  - Adoption of physician computer order entry.
  - Implementation of a unit dosing system in the hospital pharmacy.
  - Implementation of a barcoding system for medications dispensed in the hospital pharmacy.
  - Implementation of a comprehensive medication delivery system that includes all four interventions.

- *Outcome measures* – The model was used to estimate the following measures for each intervention:
  - Number of errors for each stage of the delivery system (i.e., prescription, transcription, dispensation, administration, and total errors).
  - Rates of medication errors.

- Rates of ADEs.
- ADEs by intervention.
- Additional days of hospitalization resulting from ADEs by intervention.
- Additional hospital costs resulting from ADEs

The model simulates the four stages of a hospital medication delivery system. The results indicate that clinical information systems are potentially a cost-effective means of preventing ADEs in hospitals. The results of this study indicate that an integrated medication delivery system could save up to 1,226 days of excess hospitalization and $1.4 million in associated costs in a large tertiary care hospital.

15.4.3 An Agent-based Simulation Designed to Model Events in Hospital Patient Transfers that may lead to Adverse Events

Dunn and colleagues (2011) used agent-based simulation to analyze risk associated with hospital inpatient transfers of patients. The model simulates the possible trajectories routine processes may take that deviate from prescribed work practice. The analysis helps to determine which deviations may lead to adverse events and estimates how often these deviations result in adverse events. The two adverse events that are analyzed are misidentification of a patient and compromised infection control.

- *Objective* – The aim of this study was to develop a model that can be used for risk assessment of hospital inpatient transfers.
- *Research Question #1* – Identify the variety of possible trajectories in hospital patient transfers that deviate from prescribed work practice.
- *Research Question #2* – To calculate the probability of adverse events resulting from the deviation in work practices.
- *Methods* – An agent-based simulation model was designed to represent the chain of common violations of work practices that may lead to adverse events during hospital patient transfers. Clinicians and hospital information systems were represented as interacting agents. The model simulates the inpatient transfer process using four human agents, six objects and 186 activities. Model parameters were estimated from data obtained from 101 patient transfers. Two situations were modelled: patient misidentification and violations of infection control.

- *Input Variables* – Transfers of 101 inpatients were observed. The likelihood of violations such as failure to perform patient identification checks and failure to use adequate infection control precautions were estimated from these data.

- *Outcome Measures* – Repeated simulations were run to determine the range of potential chains of events that evolve due to individual violations by interacting agents in the hospital. The likelihood of a risk of an adverse event occurring by the end of the chain of events was calculated for patient misidentification and for violations of infection control procedures.

The analysis found that 95% of simulations of patient misidentification and infection control violations were unique. This finding suggests that the process of inpatient transfer deviates from prescribed work practices in a wide variety of ways. The risk of adverse events occurring was estimated to be 8% for misidentification and 24% for violations of infection control. The value of this simulation approach over more traditional risk analysis methods is that it permits the user to quantitatively examine how individual violations of prescribed work practices combine to create risk.

15.5 Implications

The applicability of the methods described in this chapter depends upon the nature of the eHealth application, the availability of data, and the assumptions upon which the analytic approach is based. Regression analysis is used to predict one or more outcome measures based on a set of predictor variables. The purpose is to make inferences to a population from which the sample of data is drawn. The data must meet certain assumptions such as: (a) the sample of data must accurately represent the population from which it is drawn; (b) the variables are accurately measured; and (c) the relationship between the dependent variables and independent variables is correctly specified. However, there are alternative ways of estimating the equations' parameters in the event that some of these assumptions are not met.

Network analysis takes a different approach. It is used to study relationships between individuals, objects, or events, such as communication or professional ties. The nature of the relations among actors in the network may affect an actor's perceptions or actions. Data in this instance is collected on the relations among a set of actors who make up the network (e.g., a medical practice). The analysis tries to uncover significant and influential positions in the network such as opinion leaders.

Simulation involves building a dynamic model that represents a system (e.g., the emergency department of a hospital). The model involves inputs (e.g., patient arrivals) and outputs (e.g., average time to process a patient). Simulation

runs are made and the behaviour of the system is observed (e.g., crowding in the ED).

The three methods can also be used in conjunction with one another. A study of the cost-effectiveness of coronary bypass graft operations by Anderson, Harshbarger, Weng, and Anderson (2002) utilized SEM to estimate parameters of a computer simulation model. Cusumano-Towner and colleagues (2013) used social network analysis and computer simulation to study outbreaks of nosocomial infections among hospital patients.

15.6 Summary

This chapter describes three different analytic approaches to the evaluation of eHealth systems. These methods are regression analysis, network analysis, and computer simulation. Case studies are provided as examples of these approaches.

References

Anderson, J. G. (2002a). Evaluation in health informatics: Social network analysis. *Computers in Biology and Medicine, 32*(3), 179–193.

Anderson, J. G. (2002b). Evaluation in health informatics: Computer simulation. *Computers in Biology and Medicine, 32*(3), 151–164.

Anderson, J. G. (2002c). A focus on simulation in medical informatics. *Journal of the American Medical Informatics Association, 9*(5), 554–556.

Anderson, J. G., Chaturvedi, A. R., & Cibulskis, M. (2007). Simulation tools for developing policies for complex systems: Modeling the health and safety of refugee communities. *Health Care Management Science, 10*(4), 331–339.

Anderson, J. G, Harshbarger, W., Weng, W. C., & Anderson, M. M. (2002). Modeling the costs and outcomes of cardiovascular surgery. *Health Care Management Scie*nce, *5*(2), 103–111.

Anderson, J. G., Jay, S. J., Anderson, M., & Hunt, T. N. J. (2002). Evaluating the capability of information technology to prevent adverse drug events: A computer simulation approach. *Journal of the American Medical Informatics Association, 9*(5), 479–490.

Anderson, J. G., Jay, S. J., Cathcart, L., Clevenger, S. J., Perry, J., & Anderson, M. M. (1988). Physician use of clinical information systems: A computer

simulation model. In *System science in health care, vol. 1*: *Information in health care systems* (pp. 421–424). Paris: Masson.

Anderson, J. G., Ramanujam, R., Hensel, D., & Sirio, C. (2010). Reporting trends in a regional medication error data-sharing system. *Healthcare Management Science, 13*(1), 65–73.

Bardhan, I. R., & Thouin, M. F. (2013). Health information technology and its impact on the quality and cost of healthcare delivery. *Decision Support Systems, 55*(2), 438–449.

Cho, J., Park, D., & Lee, H. E. (2014). Cognitive factors of using health apps: Systematic analysis of relationships among health consciousness, health information orientation, eHealth literacy, and health app use efficacy. *Journal of Medical Internet Research, 16*(5), e125. doi: 10.2196/jmir.3283

Cusumano-Towner, M., Li, D. Y., Tuo, S., Krishnan, G., & Maslove, D. M. (2013). A social network of hospital acquired infection built from electronic medical record data. *Journal of the American Medical Informatics Association, 20*(3), 427–434.

Dunn, A. G., Ong, M. S., Westbrook, J., Magrabi, F., Coiera, E., & Wobcke, W. (2011). A simulation framework for mapping risks in clinical processes: The case of in-patient transfers. *Journal of the American Medical Informatics Association, 18*(3), 259–266.

Encinosa, W. E., & Bae, J. (2011). Health information technology and its effects on hospital costs, outcomes, and patient safety. *Inquiry, 48*(4), 288–303.

Hoot, N. R., LeBlanc, L. J., Jones, I., Scotty, R. L., Zhou, C., Gadd, C. S., & Aronsky, D. (2009). Forecasting emergency department crowding: A prospective real-time evaluation. *Journal of the American Medical Informatics Association, 16*(3), 338–345.

Kline, R. B. (2010). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford Press.

Li, N., Orrange, S., Kravitz, R. L., & Bell, R. A. (2014). Reasons for and predictors of patients' online health information seeking following a medical appointment. *Family Practice, 31*(5), 550–556. doi: 10.1093/fampra/cmu034

Mitsutake, S., Shibata, A., Ishii, K., & Oka, K. (2012). Association of eHealth literacy with colorectal cancer knowledge and screening practice among internet users in Japan. *Journal of Medical Internet Research, 14*(6), e153. doi: 10. 2196/jmir.1927

Moehr, J. R. (2002). Evaluation of health information systems: Beyond efficiency and effectiveness. *Computers in Biology and Medicine, 32*(3), 111–112.

Neter, E., & Brainin, E. (2012). eHealth literacy: Extending the digital divide to the realm of health information. *Journal of Medical Internet Research, 14*(1), e19. doi: 10.2196/jmir.1619

Zheng, K., Padman, R., Krackhardt, D., Johnson, M. P., & Diamond, H. S. (2010). Social networks and physician adoption of electronic health records: Insights from an empirical study. *Journal of the American Medical Informatics Association, 17*(3), 328–336.

Chapter 16

Methods for Data Quality Studies

Francis Lau

16.1 Introduction

The proliferation of eHealth systems has led to a dramatic increase in the volume of electronic health data being collected. Such data are often collected as part of direct patient care delivery to document the patient's conditions and the care being provided. When the collected health data are used as intended, it is referred to as primary use. Once collected the health data can be used for other purposes such as clinical quality improvement, population health surveillance, health systems planning and research. These are referred to as secondary uses (Safran et al., 2007). In Canada, a further distinction is made where all secondary uses except for research are labelled as health system use (Canadian Institute for Health Information [CIHI], 2013).

The quality of routinely collected eHealth data is a major issue for healthcare organizations. To illustrate, a systematic review by Thiru, Hassey, and Sullivan (2003) on EHR data quality in primary care found a great deal of variability depending on the type of data collected. In 10 EHR studies on sensitivities they found data completeness ranged from 93% to 100% for prescriptions, 40% to 100% for diagnoses, 37% to 97% for lifestyle in alcohol use and smoking, to 25% for socioeconomic data. A 2010 review by Chan, Fowles, and Weiner (2010) showed that the variability in the quality of EHR data is an ongoing issue especially with problem lists and medications. Iron and Manuel (2007) reported in their environmental scan of administrative health data quality assessment that: the concepts of accuracy and validity are often confused; there are no standard methods for measuring data quality; and the notion of data quality depends on the purpose for which the data are used. These findings suggest data quality can affect the performance of the eHealth systems and care delivery within organizations.

In this chapter we describe approaches to eHealth data quality assessment that are relevant to healthcare organizations. The approaches cover concepts, practice and implications. Concepts refer to dimensions, measures and methods of data quality assessment. Practice refers to how data quality assessment is done in different settings as illustrated through case examples. Implications refer to guidance and issues in eHealth data quality assessment to be considered by healthcare organizations.

16.2 Concepts of Data Quality

In this section we describe the key concepts in eHealth data quality. These are the conceptual quality dimensions, measures used to assess the quality dimensions, and methods of assessment. These concepts are described below.

16.2.1 Data Quality Dimensions

An overriding consideration when defining data quality concepts is "fitness for use". This suggests data quality is a relative construct that is dependent on the intended use of the data collected. Different terms have been used to describe the conceptual dimensions of data quality, with no agreement on which should be the standard. Sometimes the meanings of these terms overlap or conflict with each other. Drawing on the studies by Weiskopf and Weng (2013) and Bowen and Lau (2012), we arrived at the following five commonly cited terms for this chapter:

- *Correctness* – Reflects the true state of a patient, also known as accuracy. An example is whether a high blood pressure value for a patient is true or not.

- *Completeness* – Covers all truths on a patient, also known as comprehensiveness. An example is the blood pressure measurement that contains the systolic and diastolic pressures, method of assessment, and date/time of assessment.

- *Concordance* – Agreement of the data with other elements or sources, also known as reliability, consistency and comparability. An example is the use of metformin as a diabetic medication in the presence of a diabetes diagnosis.

- *Plausibility* – Does the data make sense in what is being measured given what is known from other elements? This is also known as validity, believability and trustworthiness. An example is the presence of a hypertension diagnosis in the presence of recent abnormal blood pressure measurements.

- *Currency* – Reflects the true state of a patient at a given point in time, also known as timeliness. An example is the presence of a recent blood pressure measurement when considering a hypertensive condition.

In the literature review by Weiskopf and Weng (2013), completeness, correctness and concordance were the most common dimensions assessed. Other less common data quality dimensions described in the literature (Bowen & Lau, 2012) include comprehensibility, informative sufficiency, and consistency of capture and form. These terms are defined below.

- *Comprehensibility* – The extent to which the data can be understood by the intended user.
- *Informative sufficiency* – The extent to which the data support an inference on the true state of condition.
- *Consistency of capture* – The extent to which the data can be recorded reliably without variation by users.
- *Consistency of form* – The extent to which the data can be recorded reliably in the same medium by users.

16.2.2 Data Quality Measures

The dimensions of correctness and completeness can be quantified through such measures as sensitivity, specificity, positive predictive value and negative predictive value. Quantifying these data quality measures requires some type of reference standard to compare the data under consideration. Using a health condition example such as diabetes, we can take a group of patients where the presence or absence of their condition is known, and compare with their charts to see if the condition is recorded as present or absent. For instance, if the patient is known to have diabetes and it is also recorded in his chart then the condition is true. The comparison can lead to different results as listed below.

- *Sensitivity* – The percentage of patients recorded as having the condition among those with the condition.
- *Specificity* – The percentage of patients recorded as not having the condition among those without the condition.
- *Positive predictive value* – The percentage of patients with the condition among those recorded as having the condition (i.e., condition present).

- *Negative predictive value* – The percentage of patients without the condition among those recorded as not having the condition (i.e., condition absent).

- *Correctness* – The percentage of patients with the condition among those recorded as having the condition. It can also be the percentage of patients without the condition among those recorded as not having the condition. These are also known as the positive predictive value and negative predictive value, respectively. Often only positive predict value is used to reflect correctness.

- *Completeness* – The percentage of patients recorded as having the condition among those with the condition. It can also be the percentage of patients recorded as not having the condition among those without the condition. These are also known as sensitivity and specificity, respectively. Often only sensitivity is used to reflect completeness.

The comparison of the patients' actual condition against the recorded condition can be enumerated in a 2x2 table (see Table 16.1). The actual condition represents the true state of the patient, and is also known as the reference standard.

Table 16.1
Calculation of Completeness and Correctness Using Sensitivity and Positive Predictive Value

		Reference Standard Data		
		Condition is Present	**Condition is Absent**	
Data under evaluation	Condition Appears Present	A – True Positive	B – False Positive	Correctness Positive Predictive Value (PPV) = A/(A+B) in %
	Condition Appears Absent	C – False Negative	D – True Negative	Negative Predictive Value (NPV) = D/(C+D) in %
		Completeness Sensitivity = A/(A+C) in %	Specificity = D/(B+D) in %	

Note. From"Defining and evaluating electronic medical record data quality within the Canadian context," by M. Bowen and F. Lau, 2012, *ElectronicHealthcare, 11*(1), e5–e13.

16.2.3 Data Quality Methods

Different data quality assessment methods have been described in the literature. Some methods are focused on ways to measure different dimensions of data

quality such as correctness and completeness of the data in an eHealth system. Others are concerned with the means of carrying out and reporting data quality assessment studies. There are also methods that apply predefined criteria to identify and validate specific health conditions recorded in the eHealth system. The types of methods covered in this chapter are defined below and elaborated in the next section.

- Validation of data from single and multiple sources – The use of predefined knowledge and query rules to validate the integrity of the data in one or more eHealth systems and/or databases.

- Designing, conducting and reporting data quality studies – The use of a systematic process to carry out data quality assessment studies.

- Identification and validation of health conditions – The use of predefined criteria to identify and validate specific health conditions in an eHealth system or database. The process is also known as case definition or case finding, and the criteria may be from evidence-based guidelines or empirically derived with expert consensus.

16.3 Methods of Data Quality Assessment

This section examines the three types of data quality assessment methods defined in section 16.2.3. Most of the methods were developed as part of data quality assessment studies or as validation of previously developed methods. The analysis in these methods typically involves the use of frequency distributions, cross-tabulations, descriptive statistics and comparison with a reference source for anomalies. These methods are described below.

16.3.1 Validation of Data from Single and Multiple Sources

Brown and Warmington (2002) introduced Data Quality Probe (DQP) as a method to assess the quality of encounter-driven clinical information systems. The principle behind DQP is that predefined queries can be created from clinical knowledge and guidelines to run against the system such as an EHR as measures of its quality. Typically the DQPs examine two or more data recordings that should or should not appear together in the patient record. The most common DQPs involve checking for the presence of a clinical measurement that either should always or never be associated with a diagnosis, and a therapy that either should always or never be accompanied by a diagnosis or clinical measurement. In an ideal system there should be no data inconsistencies detected when the queries are run. Examples are the absence of Hemoglobin A1c (HbA1c) test results on patients with diabetes and prescriptions of penicillin on patients with a penicillin allergy. Two types of errors can be detected: failure to record the

data or error of omission, and suboptimal clinical judgment or error of commission. Once detected, these errors should be reported, investigated and corrected in a timely fashion. To be effective, the DQPs should be run periodically with reports of any inconsistencies shared with providers at the individual and/or aggregate level for purposes of education or action.

Kahn, Raebel, Glanz, Riedlinger, and Steiner (2012) proposed a two-stage data quality assessment approach for EHR-based clinical effectiveness research that involves single and multiple study sites. In stage-1, source datasets from each site are evaluated using five types of data quality rules adapted from Maydanchik (2007). In stage-2, datasets from multiple sites are combined, with additional data quality rules applied, to compare the individual datasets with each other. Such multisite comparisons can reveal anomalies that may not be apparent when examining the datasets from one site alone. The five types of stage-1 data quality rules are outlined below (for details, see Kahn et al., 2012, p. S26, Table 3).

- Attribute domain constraints – Rules that restrict allowable values in individual data elements using assessment methods of attribute profiling, optionality, format, precision and valid values to find out-of-range, incorrect format or precision, missing or unusual data values. An example is a birthdate that is missing, unlikely, or in the wrong format.

- Relational integrity rules – Rules that ensure correct values and relationships are maintained between data elements in the same table or across different tables. An example is the use of diagnostic codes that should exist in the master reference table.

- Historical data rules – Rules that ensure correct values and relationships are maintained with data that are collected over time. For example, the recording of HbA1c results over time should correspond to the requested dates/times, follow an expected pattern, and be in a consistent format and unit.

- State-dependent objects rules – Rules that ensure correct values are maintained on data that have expected life cycle transitions. An example is a hospital discharge event should always be preceded by an admission or transfer.

- Attribute dependency rules – Rules that ensure the consistency and plausibility of related data on an entity. An example is the birthdate of a patient should not change over time; neither should a test be ordered on a deceased patient.

For stage-2 data quality rules, the focus is on semantic consistency to ensure data from different sites have the same definitions so they can be aggregated and analyzed meaningfully. The rules typically compare frequency distributions, expected event rates, time trends, missing data and descriptive statistics (e.g., mean, median, standard deviation) of the respective datasets to detect patterns of anomalies between sites. An example is the need to distinguish random versus fasting glucose tests or dramatic differences in the prevalence of diabetes between sites.

In both stage-1 and stage-2 data quality assessment, documentation is needed to articulate the rationale, methods and results of the assessments done. Often there can be hundreds of data quality rules depending on the complexity of the eHealth system and databases involved. Therefore some type of prioritization is needed on the key data elements and assessments to be included. The outputs generated can be daunting especially if every error encountered on each record is reported. The detailed errors should be grouped into categories and summarized into key areas with select performance indicators to report on the overall quality of the system or database, such as the percentage of records that pass all the data quality tests.

16.3.2 Designing, Conducting and Reporting Data Quality Studies

Bowen and Lau (2012, p. e10) published a 10-step method for conducting a context-sensitive data quality evaluation. These steps provide a systematic means of planning, conducting and reporting a data quality evaluation study that takes into account the intended use of the data and the organizational context. The 10 steps are:

- Focus on an activity that requires the use of the data being evaluated.
- Determine the context in which the activity is carried out, including the intent, tasks, people and results.
- Identify the tools/resources needed to evaluate the quality of the data and their alignment with the activity.
- Determine the degree of fitness between the activity and data being evaluated and the acceptable level of fitness.
- Select an appropriate data quality measurement method for the chosen fitness dimension being evaluated.
- Adapt the measurement method depending on the context and use data quality probes to aid in the evaluation.

- Apply the tools/resources identified in step-3 to evaluate the quality of the data being measured.
- Document the output of the fitness evaluation for each data being measured.
- Describe the overall fitness of the important data with the activity in context.
- Present data quality evaluation findings and provide feedback on the quality/utility of the data and improvement.

16.3.3 Identification and Validation of Health Conditions

Wright and colleagues (2011, pp. 2–6) developed and validated a set of rules for inferring 17 patient problems from medications, laboratory results, billing codes and vital signs found in the EHR system. A six-step process was used to develop and validate the rules. Additional analyses were done to adjust rule performance based on different rule options. These steps are listed below:

- Automated identification of problem associations – based on a previous study that identified lab-problem and medication-problem associations against gold standard clinical references and significant co-occurring statistics.
- Selection of problems of interest – an initial list of problems ranked according to three criteria: related pay-for-performance initiatives; existing decision support rules in EHR; strength of identified associations.
- Development of initial rules – confirmed relevant lab tests, medications and billing codes with medical references, added relevant free-text entries, then drafted initial rules which were reviewed by expert clinicians.
- Characterization of initial rules and alternatives – focused on patients with at least one relevant medication, lab test, billing code and vital sign but without the problem recorded in EHR; applied initial rules to identify rule-positive and rule-negative patients, then conducted chart review on a sample of patients to see if they have the problem (i.e., chart-positive, chart-negative); derived sensitivities, specificities, Positive Predictive Value (PPV) and Negative Predictive Value (NPV) of initial rule options by varying their thresholds such as lab values, drugs and counts.

- Selection of the final rule – had expert clinicians review different rule options for each problem with varying sensitivities, specificities, PPV and NPV; selected final rules with high PPV over specificity.

- Validation of the final rule – repeated above steps using an independent patient dataset from the same population.

- Additional analyses – derived sensitivity, specificity, PPV and NPV with coded problems then billing data only to adjust final set of rules based on F-measure for higher PPV over sensitivity (false negatives versus false positives).

An example of the final rules for diabetes is shown below (Wright et al., 2011, supplementary data):

- Rule 0: code or free-text problem on problem list for diabetes.

- Rule 1: any HbA1c result greater than or equal to 7.

- Rule 2: 2 or more ICD-9 billing codes in diabetes (250, 250.0, 250.00, 250.01, 250.02, 250.03, 250.1, 250.10, 250.11, 250.12, 250.13, 250.2, 250.20, 250.21, 250.22, 250.23, 250.3, 250.30, 250.31, 250.32, 250.33, 250.4, 250.41, 250.42, 250.43, 250.5, 250.50, 250.51, 250.52, 250.53, 250.6, 250.60, 250.61, 250.62, 250.63, 250.7, 250.71, 250.72, 250.73, 250.8, 250.80, 250.81, 250.82, 250.83, 250.9, 250.91, 250.92, 250.93).

- Rule 3: at least one medication in injectable anti-diabetic agents or oral anti-diabetic agents.

16.4 Case Examples

This section includes two published examples of eHealth data quality studies: one is on multisite data quality assessment while the other is on primary care EMRs.

16.4.1 Multisite Data Quality Assessment

Brown, Kahn, and Toh (2013) reviewed multisite data quality checking approaches that have been field-tested in distributed networks for comparative effectiveness research such as the Observational Medical Outcomes Partnership in the United States. Typically these networks employ a common data model and different types or levels of data quality checks for cross-site analysis as described in Kahn's 2-stage data quality assessment approach (Kahn et al., 2012). These data quality-checking approaches are:

- Common data model adherence – These are checks on extracted data against the common data model dictionary for consistency and adherence to the model. They are: (a) syntactic correctness on transformed variable names, values, lengths and format meeting data model specifications; (b) table structure and row definition correctness; and (c) cross-table variable relationships for consistency. Examples are valid codes for sex and diagnosis, linkable tables by person or encounter identifier, and presence of valid enrolment for all prescription records.

- Data domain review – These are checks on the frequency and proportion of categorical variables, distribution and extreme values for continuous variables, missing and out-of-range values, expected relationships between variables, normalized rates and temporal trends. The domains may cover enrolment, demographics, medication dispensing, prescribing, medication utilization, laboratory results and vital signs. Examples of checks are enrolment periods per member, age/sex distribution, dispensing/prescriptions per user per month, diagnoses/procedures per encounter, weight differences between men and women, and number of tests conducted per month.

- Review of expected clinical relationships with respect to anomalies, errors and plausibility – Within and cross-site co-occurrence of specific clinical variables should be assessed, such as the rate of hip fractures in 60- to 65-year-old females, male pregnancy and female prostate cancer.

- Member/study-specific checks – These are checks to ensure proprietary and privacy-related policies and regulations for specific members are protected, such as the inclusion of unique product formulary status, clinical/procedure codes, patient level information and tables with low frequency counts; and to detect data variability across study sites such as the exposure, outcome and covariates under investigation.

16.4.2 Improving Data Quality in Primary Care EMRs

The Canadian Primary Care Sentinel Surveillance Network (CPCSSN) is a pan-Canadian practice-based research network made up of 10 regional networks involving more than 1,000 primary health care providers in eight provinces and territories (CPCSSN, n.d.). Its mission is to improve primary health care delivery and outcomes, epidemiological surveillance, research excellence, and knowledge translation. The effort involves the extraction and use of EMR data from community-based primary health care practices to inform and improve the manage-

ment of the most common chronic diseases in Canada. Here we describe the CPCSSN Data Presentation Tool (DPT) that has been developed to improve the management of individual and groups of patients within and across practices (Moeinedin & Greiver, 2013). In particular, we emphasize the effort undertaken to improve the quality of the EMR data and its impact in the DPT initiative.

- DPT purpose and features – The DPT is an interactive software developed as a quality dashboard to generate custom reports at the provider, office and organizational levels. It uses EMR data that have been de-identified, cleaned and standardized through a systematic process which are then returned to the providers for use in quality improvement purposes. These include the ability to improve data quality at the practice level, re-identify at-risk patients for tracking and follow-up, and produce custom reports such as prescribing patterns and comorbidities in specific chronic diseases (Williamson, Natarajan, Barber, Jackson, & Greiver, 2013; Moeinedin & Greiver, 2013).

- DPT study design – The DPT was implemented and evaluated as a quality improvement study in a family health team in Ontario. The study used mixed methods to examine practice change before and after DPT implementation from May to August 2013. Sixty-one primary care providers took part in the study. The qualitative component included field notes, observations, key informant interviews and a survey. The quantitative component measured the change in data quality during that period {Moeinedin & Greiver, 2013; Greiver et al., 2015).

- Data quality tasks – CPCSSN has developed an automated approach to cleaning EMR data. The data cleaning algorithms are used to identify missing data, correct erroneous entries, de-identify patients, and standardize terms. In particular, the standardization process can reduce the various ways of describing the same item into one term only. Examples are the use of kilograms for weights, one term only for HA1c, and three terms only for smoking status (i.e., current smoker, ex-smoker, never smoked). The cleaned data are then returned to the providers, allowing them to assess the data cleaning needed at the local level within their EMRs. To ensure transparency, CPCSSN has published its data cleaning algorithms in peer-reviewed journals and on its website (Greiver et al., 2012; Keshavjee et al., 2014).

- Key findings – The family health team in the DPT study was able to identify missing and non-standardized data in its EMRs. The

DPT was used to produce quality reports such as the prevalence of hypertension and dementia in the region, re-identification of high-risk patients for follow-up, and specific medication recall. Overall, the updating and standardization of the EMR data led to a 22% improvement in the coding of five chronic conditions and the creation of registries for these conditions (Moeinedin & Greiver, 2013; Greiver et al., 2015).

16.5 Implications

As healthcare organizations become more dependent on eHealth systems for their day-to-day operations, the issue of eHealth data quality becomes even more prominent for the providers, administrators and patients involved. The consequence of poor-quality data can be catastrophic especially if the care provided is based on incomplete, inaccurate, inaccessible or outdated information from the eHealth systems. The data quality assessment approaches described in this chapter are empirically-derived pragmatic ways for organizations to improve the quality and performance of their eHealth systems. To do so, there are a number of policy and practice implications to be considered.

For policy implications, healthcare organizations need to be aware of the task-dependent nature of data quality, or fitness for use, in order to embark on data quality policies that are most appropriate for their needs. An important first step is to adopt a consistent set of eHealth data quality concepts with clearly defined evaluation dimensions, measures and methods. More importantly, it should be recognized that data quality evaluation is only a means to an end. Once the state of eHealth data quality has been identified, there must be remedial actions with engaged owners and users of the data to rectify the situation. Last, organizational leaders should foster a data quality culture that is based on established best practices.

For practice implications, healthcare organizations need to dedicate sufficient resources with the right expertise to tackle data quality as a routine practice. Data quality evaluation is a tedious endeavour requiring attention to detail that includes meticulous investigation into the root causes of the data quality issues identified. There should be detailed documentation on all of the data quality issues found and remedial actions taken to provide a clear audit trail for references. Last, since providers are responsible for a substantial portion of the routine clinical data being collected, they need to be convinced of the value in having high-quality data as part of patient care delivery.

16.6 Summary

This chapter described eHealth data quality assessment approaches in terms of the key concepts involved, which are the data quality assessment dimensions, measures and methods used and reported. Also included in this chapter are two

examples of data quality assessment studies in different settings, and related implications for healthcare organizations.

References

Bowen, M., & Lau, F. (2012). Defining and evaluating electronic medical record data quality within the Canadian context. *ElectronicHealthcare, 11*(1), e5–e13.

Brown, J., & Warmington, V. (2002). Data quality probes – exploiting and improving the quality of electronic patient record data and patient care. *International Journal of Medical Informatics, 68*(1–3), 91–98.

Brown, J., Kahn, M., & Toh, S. (2013). Data quality assessment for comparative effectiveness research in distributed data networks. *Medical Care, 51*(8 suppl 3), S22–S29.

Canadian Institute for Health Information (CIHI). (2013). *Better information for improved health: A vision for health system use of data in Canada.* Ottawa, ON: Author.

Chan, K. S., Fowles, J. B., & Weiner, J. P. (2010). Electronic health records and the reliability and validity of quality measures: A review of the literature. *Medical Care Research and Review, 67*(5), 503–527.

Canadian Primary Care Sentinel Surveillance Network (CPCSSN). (n.d.). *Canadian primary care sentinel surveillance network* (website). Retrieved from http://cpcssn.ca/

Greiver, M., Drummond, N., Birwhistle, R., Queenan, J., Lambert-Lanning, A., & Jackson, D. (2015). Using EMRs to fuel quality improvement. *Canadian Family Physician, 61*(1), 92.

Greiver, M., Keshavjee, K., Jackson, D., Forst, B., Martin, K., & Aliarzadeh, B. (2012). Sentinel feedback: Path to meaningful use of EMRs. *Canadian Family Physician, 58*(10), 1168.

Kahn, M. G., Raebel, M. A., Glanz, J. M., Riedlinger, K., & Steiner, J. F. (2012). A pragmatic framework for single-site and multisite data quality assessment in electronic health record-based clinical research. *Medical Care, 50*(7), S21–S28.

Keshavjee, K., Williamson, T., Martin, K., Truant, R., Aliarzadeh, B., Ghany, A., & Greiver, M. (2014). Getting to usable EMR data. *Canadian Family Physician, 60*(4), 392.

Iron, K., & Manuel, D. G. (2007, July). *Quality assessment of administrative data (Quaad): An opportunity for enhancing Ontario's health data* (ICES Investigative Report). Toronto: Institute for Clinical Evaluative Sciences.

Maydanchik, A. (2007). On hunting mammoths and measuring data quality. *Data Management Review, 6*, 14–15, 41.

Moeinedin, M., & Greiver, M. (2013). *Implementation of the data presentation tool in a primary care organization*. Presentation to Trillium Research Day, Toronto, June 19, 2013. Retrieved from http://www.trilliumresearchday.com/documents/2013_Moeinedin_DPT_implementation_Final_June192013.pdf

Safran, C., Bloomrosen, M., Hammond, W. E., Labkoff, S., Markel-Fox, S., Tang, P. C., & Detmer, D. (2007). Toward a national framework for the secondary use of health data: An American Medical Informatics Association white paper. *Journal of American Medical Informatics Association, 14*(1), 1–9.

Thiru, K., Hassey, A., & Sullivan, F. (2003). Systematic review of scope and quality of electronic patient record data in primary care. *British Medical Journal, 326*(7398), 1070–1072.

Weiskopf, N. G., & Weng, C. (2013). Methods and dimensions of electronic health record data quality assessment: Enabling reuse for clinical research. *Journal of American Medical Informatics Association, 20*(1), 144–151.

Williamson, T., Natarajan, N., Barber, D., Jackson, D., & Greiver, M. (2013). Caring for the whole practice — the future of primary care. *Canadian Family Physician, 59*(7), 800.

Wright, A., Pang, J., Feblowitz, J. C., Maloney, F. L., Wilcox, A. R., Ramelson, H. Z., Schneider, L. I., & Bates, D. W. (2011). A method and knowledge base for automated inference of patient problems from structured data in an electronic medical record. *Journal of American Medical Informatics Association, 18*(6), 859–867.

Chapter 17

Engaging in eHealth Evaluation Studies

Craig Kuziemsky, Francis Lau

17.1 Introduction

Healthcare systems worldwide are undergoing substantial transformation to enable delivery of patient-centred, safe, collaborative care. Health information technology (HIT) will play a substantial role in these transformative efforts. However, the transformation of healthcare delivery makes HIT evaluation complex as it creates a multidimensional spectrum by which HIT needs to be evaluated. For example, Bates (2015) calls coordinated care delivery the next great opportunity for informatics. In that context then, HIT needs to be evaluated based upon how well it supports care coordination. While HIT has in the past often been evaluated in a broad sense to examine the adoption of a specific task (e.g., order entry, decision support), we now recognize the need to evaluate HIT from a more holistic perspective. While HIT may be implemented to support care delivery processes in one hospital, the impact and evaluation of the system may go far beyond that hospital and include care processes in other hospitals or in the community at large.

This chapter provides a perspective on eHealth evaluation within the context of the evolving healthcare delivery system. It provides practical insight such as linking eHealth evaluation to frameworks for healthcare transformation, insight on engaging practitioners in eHealth evaluation, and ways to conduct evidence-based eHealth evaluation.

17.2 Conducting eHealth Evaluation Studies

The evaluation of eHealth has grown in complexity because there has been a significant shift in how HIT is governed. In its early years, HIT was implemented and evaluated within the boundaries of individual institutions. In fact, many of

such historic HIT systems as the HELP system (Pryor, 1988), the Regenstrief Medical Record System (McDonald et al., 1999), and Brigham Integrated Computing System (Tiech et al., 1999), were developed and maintained in-house. Over the years, in-house development gave way to large-scale vendors, leading to the current era of HIT integration beyond such traditional boundaries as hospitals and clinics and into the community and patients' homes.

This movement is in response to national governmental initiatives for designing integrated care delivery systems. Examples include Canada Health Infoway in Canada, the Connecting for Health Initiative in the United Kingdom (Hamblin & Ganesh, 2007; McGlynn, Shekelle, & Hussey, 2008) and the Health Information Technology for Economic and Clinical Health (HITECH) Act in the United States (Blumenthal, 2011). These national initiatives have shifted the landscape of HIT evaluation in that they have brought with them new expectations of the role that HIT will play. While it is always necessary to evaluate HIT from the perspective of the front-line users, national initiatives have added requirements pertaining to the demonstration of macro-level measures such as accountability, service delivery and care coordination. These must be reported on due to the desire of those who are responsible for funding and coordination levels to be more accountable for care delivery. However, these national initiatives have not gone without criticism. Canada Health Infoway and the HITECH Act have encountered difficulties achieving their objectives (Mennemeyer, Menachemi, Rahurkar, & Ford, 2015; Rozenblum et al., 2011), while mounting criticism and budget overruns led to the disbandment of the Connecting for Health Initiative in 2013.

In conducting evaluation studies we must remember that there is often a gap between HIT implementation and how it supports care delivery (Novak, Brooks, Gadd, Anders, & Lorenzi, 2012). HIT evaluation can be broadly classified into two main categories. First, is the evaluation needed to support delivery from line user interactions with HIT (i.e., the micro level); these evaluation methods were detailed in chapter 8. Second are evaluation approaches to see how well HIT supports broader care delivery objectives (i.e., the macro level). Examples of such approaches include evaluation of continuity of care or collaborative care delivery.

While micro-level evaluations have been the predominant evaluation category to date, we are seeing an increasing desire for macro-level evaluations. The Triple Aim is an example of a macro-level framework that that has been used to evaluate HIT implementation (Sheikh, Sood, & Bates, 2015). The Triple Aim has three goals: first, improving the quality, safety, and experience of care; second, enhancing population health; and third, reducing per capita costs of healthcare (Berwick, Nolan, & Whittington, 2008). However, while the HITECH Act has improved the uptake of HIT, its ability to bring about more substantial healthcare transformation (e.g., the Triple Aim) has been hampered by such factors as usability, interoperability and inappropriate funding models, for example, fee for service (Sheikh et al., 2015).

When evaluating macro-level outcomes we must ensure that a favourable macro-level outcome is not hiding implementation issues at the micro level. For

example, wait times and system throughput are common macro-level measures and thus are used as metrics for HIT evaluation. A U.K. study on national targets for emergency department wait times described how achieving a four-hour ED wait time target led to micro-level issues between physicians and patients and colleagues (Vezyridis & Timmons, 2014). Again, successfully achieving an evaluation metric at one level may come at a price of causing unintended consequences at other levels, which emphasizes the need for multilevel evaluations that look at a range of outcomes, for instance organizational, social, clinical, and cognitive (Bloomrosen et al., 2011; Kuziemsky & Peyton, 2016).

Therefore the first step in conducting eHealth evaluation is to understand the scope of evaluation at all levels and then put in place an appropriate evaluation design.

17.2.1 Good eHealth Evaluation Practices

Frameworks for conducting eHealth evaluation exist at both the micro and macro levels. Many of the previous chapters in this handbook have described frameworks at both micro (i.e., clinical) and macro (i.e., organizational and public health) levels for conducting HIT evaluation. Evidence-based evaluation approaches should be used whenever possible to ensure evaluation rigour but also to enable comparability across studies.

In chapter 8 we introduced the GEP-HI guidelines, intended to provide a set of structured principles to design and carry out evaluation studies in different IT contexts (Nykänen et al., 2011). The GEP-HI principles contains six phases that provide a practical set of considerations for how to plan, implement and execute an eHealth evaluation study. Phase one, preliminary outline, describes the purpose of the study and how the evaluation should take place. Phase two is the study design where the actual evaluation design is conceived. Phase three is the operationalization phase where the methods for the evaluation study are formalized in the context of the HIT being studied, its organizational setting and the information that is needed. Phase four is project planning where plans and procedures are developed for the evaluation study. Phase five is the actual execution of the evaluation study. Phase six is the reporting of the study results, completion of any remaining issues and closure of the study (Nykänen et al., 2011). Each of the phases has a subset of procedures that are carried out as part of each phase. For example, in phase two (study design) it is necessary to look at factors such as the project timeline, budget, ethical and legal issues, the evaluation issues and questions, and the different methods that can be used to study them. Each GEP-HI phase and accompanying items serve to structure the stages and components of an evaluation study.

17.2.2 Rapid eHealth Evaluation

Chapter 8 described how HIT evaluation must be done in a holistic manner that spans the entire system development life cycle (SDLC), from requirements elicitation to systems design and implementation.

Evaluation needs to begin as soon as requirements are elicited, continue through to model development, and finally to implementation of the HIT. Both formative and summative evaluations need to be done (McGowan, Cusack, & Poon, 2008). However, this does not mean that all evaluation studies need to go through the entire spectrum of the SDLC at both formative and summative levels. For example, if an organization already has an existing HIT in place they may need to proceed directly to do a summative evaluation of the system. Other organizations may need to start with a formative evaluation and then proceed to a summative one, depending on the level of maturity of the HIT. Regardless of the stage and type of evaluation that is done, practitioners need to be involved in HIT evaluation. Practitioners and other front-line users (e.g., managers) are the best people to provide insight on various contexts of use between HIT and work practices. Involving front-line users in HIT evaluation studies can facilitate better adoption and safer use of HIT as a way of mitigating unintended consequences from HIT implementation (Novak et al., 2012).

17.2.3 Practical Considerations

Healthcare delivery is context-dependent, which needs to be considered in any eHealth evaluation study. Evaluating a system without due consideration of context will be problematic. As described above, HIT evaluation has both micro and macro aspects to it that must be considered wherever possible. However, considering these two dimensions can often pose challenges to HIT evaluation. A consequence of this multidimensionality is that HIT evaluation may have conflicting requirements (Kuziemsky & Peyton, 2016). For example, administrators are facing increased pressure to be accountable for care delivery and the quality of services provided. Timely reporting of these outcomes necessitates the collection of data, which can pose a burden to front-line clinicians (Kuziemsky & Peyton, 2016). Therefore evaluating HIT from administrative and clinical perspectives may have different evaluation objectives. Another practical consideration is the need for upstream impacts to be measured. While HIT evaluation has historically focused on tracking services or processes in the moment — for example, how well a system facilitates order entry or tracks a patient through the emergency department — it has been emphasized that healthcare is about promoting and maintaining health, not just making services available (Butler, 2016). To that end, we need to consider upstream impacts of HIT use such as how it changes consumer behaviour as part of the developing of healthier lifestyles. This makes HIT evaluation that much more complex as the evaluation parameters may need to evolve over time. While evaluation of access to services may be an appropriate evaluation today, in the future we will be interested in how that access leads to upstream impacts such as connectivity between acute and community settings and patient engagement in care monitoring and delivery.

A key consideration is that many of the processes that HIT is automating are evolving or immature (Kuziemsky, 2016). Common health system objectives such as collaborative care delivery or patient-centred care are evolving processes

and thus evaluation metrics will need to evolve too. Healthcare systems are learning systems and therefore it is essential that system objectives be evaluated in an iterative manner (Friedman et al., 2015).

We also need to acknowledge that just because there may be a lack of evaluation evidence or an abundance of studies highlighting conflicting or adverse outcomes from HIT about HIT (Chaudry et al., 2006; Karsh, Weinger, Abbott, & Wears, 2010), it does necessarily mean all HIT is ineffective (Koppel, 2013). HIT may indeed provide benefits at patient, administration and population levels, but the complexity of the healthcare domain makes evaluation very challenging. Classic evaluation approaches, such as the randomized controlled trial, cannot be applied to HIT evaluation because of the complex reality of healthcare delivery (Koppel, 2013). HIT implementation may give completely different results in two different settings (Niazkhani, van der Sijs, Pirnejad, Redekop, & Aarts, 2009). The key message is that evaluation must strike a balance between methodological rigour and different types of evaluation methods, in light of the aforementioned need to consider formative and summative evaluation processes.

A final practical consideration is the extent of the user base that will be using a given HIT. Delivery modes such as collaborative team-based care delivery occur across multiple providers, and individuals may change work practices as part of working collaboratively (Sherer, Meyerhoefer, Sheinberg, & Levick, 2015). If HIT is meant to support team-based care delivery, then it must be evaluated from the perspective of the different team members who will be using the system (Kuziemsky & Kushniruk, 2014).

17.3 Reporting of eHealth Evaluation Studies

Further to the above point about the need for better evidence on how and why HIT works in different circumstances is the need for common reporting of HIT evaluation studies to enable comparison across settings. To that end, there has been the development of guidelines to enable consistent reporting of HIT evaluation. The statement on reporting of evaluation studies in health informatics (STARE-HI) guidelines, first introduced in chapter 8, is one such example. This chapter describes STARE-HI in more detail.

17.3.1 STARE-HI Guidelines

The STARE-HI guidelines were first established in 2009 to provide consistency in how an HIT evaluation study is reported as part of improving the evidence base of health informatics evaluations (Talmon et al., 2009). The overarching goal of STARE-HI is to enable a reader to determine whether or not the design, the outcome and the derived conclusions of an HIT evaluation study are valid (Brender et al., 2013).

STARE-HI contains 35 items to frame how an HIT evaluation study is reported from the formulation of title and abstract to the description of the study context, objectives and methods, results and conclusion (Talmon et al., 2009). Each sec-

tion then has specific details that should be included in the report. For example, the methods section should include details on the study design, theoretical background, participants, study flow, outcome measures or evaluation criteria, methods for data acquisition and measurement, and methods for data analysis (Talmon et al., 2009). The study context section of STARE-HI is particularly important for helping the generalizability of an evaluation study. The organizational setting should be described, for example, the geographical location and type of facility where the HIT is deployed (e.g., primary, secondary, tertiary care, home care). In addition, any specifics should be listed, such as whether a system is only used in a particular unit of a setting (e.g., an intensive care unit) as well as details on the type of system (e.g., laboratory, computer provider order entry). It should be noted whether the system is designed in-house or is a commercial product and the types of tasks it supports (Talmon et al., 2009). A comprehensive case example of using STARE-HI is provided by Brender and colleagues (2013).

Aside from providing consistency in reporting, STARE-HI also enables easier determination of which papers can be used in meta-analyses of health informatics interventions (Talmon et al., 2009). STARE-HI has been formally endorsed by the International Medical Informatics Association (IMIA). While the overall goal of STARE-HI is to develop standards for how HIT evaluation studies are reported, the developers of STARE-HI emphasize that it is meant to be used as a guideline, not a prescriptive structural standard (Talmon et al., 2009; Brender et al., 2013). The manner in which an HIT evaluation study is described and the degree of detail on each item will vary from study to study and may be influenced by the requirements of the journal where the study is being published (Talmon et al., 2009). Further, not all issues are relevant to every study and HIT evaluators need to consider which of the guidelines and recommendations are valid for a particular HIT evaluation context (Brender et al., 2013).

17.3.2 Mini-STARE-HI Guidelines

An acknowledged shortcoming with STARE-HI is that it relies on journal articles while ignoring the wide knowledge base contained in conference proceedings. To address that issue, mini STARE-HI guidelines were developed to guide authors in using the STARE-HI guidelines for a conference paper (de Keizer et al., 2010).

17.4 eHealth Evaluation Resources

A number of resources exist to help guide eHealth evaluation practices. A few of these resources are described below.

17.4.1 UVic eHealth Observatory

The University of Victoria (UVic) eHealth Observatory in British Columbia, Canada, is an example of a grant-funded research program to engage the eHealth community in advancing the science and practice of eHealth evaluation through knowledge creation and translation, and capacity building. It was part

of a five-year eHealth Chair program that was jointly funded by the Canadian Institutes for Health Research and Canada Health Infoway. The overall aim of the Observatory was to monitor the effects of eHealth system deployment in Canada. The specific objectives were to: (a) employ rigorous models, methods and metrics to evaluate eHealth system adoption/use and impact; (b) engage the eHealth community in knowledge translation (KT) to synthesize, share, and use the knowledge gained; and (c) build research capacity in eHealth system implementation and evaluation through graduate education and training. There were three program components:

- *Research Innovation* – This component was to: (a) consolidate existing evidence on eHealth evaluation models, methods and metrics; (b) apply rapid methods to evaluate eHealth system adoption/use and impact; (c) apply rapid methods to evaluate secondary use of eHealth data in performance management.

- *Mentoring/Education* – This component was to build eHealth evaluation research capacity by establishing a research/training environment and learning modules for educational programs and professional development.

- *Linkage/Exchange* – This component focused on integrated KT by engaging potential knowledge users in the entire eHealth evaluation research process. It covered setting the questions, deciding on the methodology, being involved in data collection and tools development, interpreting the findings, and disseminating results.

Over the five-year period, the UVic eHealth Observatory has had tangible impacts in advancing the science and practice of eHealth evaluation in Canada and elsewhere. Examples of the outputs include:[1]

- *Expanded Evidence Base* – Contribution to the growing eHealth evaluation evidence base in the form of: (a) systematic reviews on the current state of evidence on eHealth systems, physician office EMRs, medication reconciliation and economic evaluation; (b) field evaluation studies on the impacts of primary and ambulatory care EMRs; (c) use of palliative performance scale to provide meaningful survival estimates; and (d) primary and secondary use of SNOMED CT in primary and palliative care.

1 UVic eHealth Observatory. URL: http://ehealth.uvic.ca/index.php

- *Conceptual Frameworks* – Four frameworks have been developed as mental models to make sense of eHealth under different contexts. They are the: (a) Clinical Adoption Framework that was built on the micro level Benefits Evaluation Framework expanded to include the meso organizational level and the macro societal level; (b) Clinical Adoption Meta-Model that describes how evaluation should evolve over the life cycle of eHealth adoption; (c) Economic Evaluation Model that describes the key components of eHealth economic evaluation design; and (d) eHealth Value Framework that describes the dynamic interactions among eHealth investment, adoption and value.

- *Pragmatic Methodologies* – eHealth implementation and evaluation methods that have been developed include: (a) rapid evaluation methods for conducting field EMR evaluation studies; (b) encoding and evaluation methods for SNOMED CT; (c) Web-based surveillance tools for palliative end-of-life care with existing eHealth data sources; and (d) a technical report and an inventory of eHealth benefits evaluation methods and metrics.

- *Virtual Learning Communities* – A virtual community of over 100 eHealth practitioners and researchers has been created to take part in an ongoing monthly series of webinar sessions on a variety of topics related to eHealth evaluation. Participants also had opportunities to share ideas and lessons from their own implementation and evaluation experiences within their organizations.

- *Highly Qualified Personnel* – Close to 50 individuals have received eHealth evaluation-related education/training. They included trainees pursuing undergraduate and graduate health informatics degrees at UVic, as well as postdoctoral fellows, practising clinicians and research analysts working on evaluation-related projects funded by the Observatory and collaborating partners.

17.4.2 Infoway's Benefits Evaluation Program

The Benefits Evaluation (BE) strategy[2] at Canada Health Infoway is one example of the effort made at the national level to engage stakeholder organizations across Canada in making eHealth evaluation a part of their eHealth strategy. Infoway is an independent non-profit corporation funded by the Canadian federal and provincial governments to accelerate the development, adoption and use of dig-

2 Infoway Benefits Evaluation Framework and Strategy. URL: https://www.infoway-inforoute.ca/en/solutions/benefits-evaluation/benefits-evaluation-framework

ital health across the country. The overall goal of Infoway's BE strategy is to help understand the impacts of eHealth solutions on individuals, organizations and the healthcare system as a whole. The BE strategy has several components:

- *BE Framework* – Infoway has worked with a panel of researchers to develop the BE Framework (see chapter 2) as a conceptual model to describe the relationship between the adoption of an eHealth solution and its effects. While such contextual factors as organizational strategy, culture and process are considered out of scope, the BE Framework provides a useful organizing scheme to understand and measure the effects, identify the barriers and communicate the successes of eHealth adoption. Since its creation, the BE Framework has been applied across Canada and internationally to eHealth investments to evaluate their benefits and guide future initiatives.

- *Change Management Framework* – Infoway has also recommended the integration of BE with its National Change Management (CM) Framework, which has been developed to describe the change management activities needed when adopting eHealth solutions. The framework has six core elements: governance and leadership; stakeholder engagement; communications; workflow analysis and integration; training and education; and monitoring and evaluation. Collectively, the BE and CM Frameworks represent the current state of best practices in helping to achieve tangible values from the adoption of eHealth solutions.

- *BE Indicators Technical Report Version 2.0* – This report contains an inventory of empirical BE methods, measures and tools for different eHealth domains such as imaging, lab and drug information systems, interoperable EHR viewers, EMRs, telehealth, consumer health, and public health surveillance. It also contains summaries of completed BE studies and lessons learned from jurisdictional eHealth systems adopted across the country.

- *BE Resource Inventory* – These are resources assembled by Infoway to support jurisdictions in implementing, adopting and evaluating their eHealth solutions. They include the BE and CM Frameworks, the BE Indicators Technical Report, various BE methods and tools, jurisdictional BE reports and BE-related publications. Examples include the Infoway System and Use Assessment survey instrument for measuring eHealth system use and satisfaction, the BE report on Emerging Benefits of Ambulatory Care EMRs in Canada, as well as the CM Toolkit that is made up of assessment templates, work-

flow analysis checklist and sample evaluation methods. A guidance document has also been published by Infoway on the principles for sharing methods and data, as well as communicating results.

- *Pan-Canadian BE and CM Networks* – Infoway has established the BE and CM Networks to promote the sharing of best practices, the communication of BE study findings and lessons, in addition to contributing to the development of BE indicators among its network members. They include jurisdictional eHealth team leaders and members, eHealth practitioners from healthcare organizations, and eHealth researchers from research/academic institutions. Periodic face-to-face and virtual meetings and online discussion forums are held to facilitate these networking activities.

17.4.3 Other Useful Resources

Austria's University for Health Sciences, Medical Informatics and Technology (UMIT) has an inventory of eHealth evaluation publications, compiled by Professor Dr. Elske Ammenwerth, that can be searched using various criteria including language, type of system (e.g., EHR, CPOE), country of origin, and type of evaluation study.

Another resource is the Agency for Healthcare Research and Quality (AHRQ) of the United States Department of Health and Human Services, which offers numerous resources for patients, professionals and policy-makers. Resources specific to evaluation include a health IT evaluation toolkit and set of evaluation measures, quick reference guides, a toolkit for workflow assessment for health IT and a toolkit for human factors design for consumer Health IT in the home.

A number of other eHealth evaluation resources exist, including resources from organizations such as the International Medical Informatics Association, the American Medical Informatics Association and the Healthcare Information and Management Systems Society (HIMSS). Country-specific resources also exist, such as the aforementioned Canada Health Infoway and the Office of the National Coordinator for Health Information Technology in the United States.

17.5 Summary

This chapter expands upon some of the content from previous chapters by providing practical insight for conducting eHealth evaluation studies. It emphasized the relationship between macro-level healthcare system delivery initiatives and the micro level where care delivery is actually provided. Governments throughout the world are relying upon HIT to help transform healthcare delivery into integrated patient-centred care delivery systems that support care delivery across providers and settings.

Examples of such healthcare transformation initiatives include the Triple Aim and Accountable Care Initiatives from the United States, and Canada

Health Infoway in Canada. While HIT may indeed be a key driver of healthcare transformation, a key aspect of HIT evaluation is to understand how macro-level transformation initiatives may impact care delivery at the micro level. Measuring such macro-level outcomes as access to services or care integration across settings can lead to unintended consequences issues, for example workflow or communication issues at the micro level.

A key challenge in reconciling the micro and the macro is that priorities may differ across micro and macro levels. Governments and health authorities often want to collect data to track patient access to services or wait times for services, but the burden to collect the data falls on front-line clinicians (Kuziemsky & Peyton, 2016). These different priorities put an increased emphasis on the need to involve practitioners at all levels of eHealth evaluation in order to understand both the "in-the-moment" and upstream implications of HIT.

This longitudinal evaluation approach is a significant shift from how HIT evaluation used to be done where it largely focused on the technology itself. While Health IT and the broader IT community have made significant progress in developing models and frameworks for studying user interactions with HIT (e.g., the Technology Adoption Model), and usability and cognition evaluation, the erosion of the boundaries between micro, meso and macro systems require us to evaluate HIT beyond the day-to-day usage.

We also need to strive towards developing more evidence around HIT evaluation. With respect to evidence-based HIT evaluation, the point made by Koppel (2013) needs to be emphasized — that just because there is a shortcoming of evidence on HIT, it does not mean that HIT does not work. Rather, the complexity and multiple contexts within which healthcare delivery takes place makes it very difficult to develop evidence that is applicable across all settings. We therefore need to continue to research healthcare complexity and contexts to guide HIT evaluation. We also need to recognize that healthcare systems are learning systems and, thus, processes. Therefore there is a need to evaluate them from the context of the evolution of processes (Friedman et al., 2015).

A significant challenge in eHealth evaluation is the need for comparability across settings. Relationship building with the practitioners is a significant part of HIT evaluation. This chapter described two evaluation guidelines (GEP-HI and STARE-HI), which are used, respectively, for conducting and reporting HIT evaluation studies. It is essential for practitioners to be involved in HIT evaluation and GEP-HI provides a practical set of guidelines for involving practitioners in eHealth evaluation as way of establishing relationships. This chapter also provided examples of resources for conducting HIT evaluation, again emphasizing the practical aspects of evaluation.

References

Bates, D. W. (2015). Health information technology and care coordination: The next big opportunity for informatics? *Yearbook of Medical Informatics*, *10*(1), 11.

Berwick, D. M., Nolan, T. W., & Whittington, J. (2008). The triple aim: care, health, and cost. *Health Affairs, 27*(3), 759–769.

Bloomrosen, M., Starren, J., Lorenzi, N. M., Ash, J. S., Patel, V. L., & Shortliffe, E. H. (2011). Anticipating and addressing the unintended consequences of health IT and policy: a report from the AMIA 2009 Health Policy Meeting. *Journal of the American Medical Informatics Association, 18*(1), 82–90.

Blumenthal, D. (2011). Wiring the health system — Origins and provisions of a new federal program. *New England Journal of Medicine, 365*(24), 2323–2329.

Brender, J., Talmon, J., de Keizer, N., Nykänen, P., Rigby, M., & Ammenwerth, E. (2013). STARE-HI: Statement on reporting of evaluation studies in health informatics, explanation and elaboration. *Applied Clinical Informatics, 4*(3), 331–358.

Butler, S. M. (2016). The future of the Affordable Care Act: Reassessment and revision. *Journal of the American Medical Informatics Association, 316*(5), 495–497. doi: 10.1001/jama..9881

Chaudhry, B., Wang, J., Wu, S., Maglione, M., Mojica, W., Roth, E., Morton, S. C., & Shekelle, P. G. (2006). Systematic review: Impact of health information technology on quality, efficiency, and costs of medical care. *Annals of Internal Medicine, 144*(10), 742–752.

de Keizer, N. F., Talmon, J., Ammenwerth, E., Brender, J., Nykänen, P., & Rigby, M. (2010). Mini STARE-HI: Guidelines for reporting health informatics evaluations in conference papers. In C. Safran, S. Reti, & H. F. Marin (Eds.), *MEDINFO 2010: Proceedings of the 13th World Congress on Medical Informatics* (pp. 1206–1210). Amsterdam: IOS Press. ISBN 978-1-60750-587-7.

Friedman, C., Rubin, J., Brown, J., Buntin, M., Corn, M., Etheredge, L., … Van Houweling, D..(2015). Toward a science of learning systems: a research agenda for the high-functioning Learning Health System. *Journal of the American Medical Informatics Association, 22*(1), 43–50.

Hamblin, R., & Ganesh, J. (2007). *Measure for measure: Using outcome measures to raise standards in the NHS*. London: Policy Exchange.

Karsh, B.-T., Weinger, M. B., Abbott, P. A., & Wears, R. L. (2010). Health information technology: Fallacies and sober realities. *Journal of the American Medical Informatics Association, 17*(6), 617–623

Koppel, R. (2013). Is healthcare information technology based on evidence? *IMIA Yearbook of Medical Informatics, 8*, 7–12.

Kuziemsky, C. E. (2016). Decision-making in healthcare as a complex adaptive system. *Healthcare Management Forum, 29*(1), 4–7.

Kuziemsky, C. E., & Kushniruk, A. (2014). Context mediated usability testing. In C. Lovis, B. Seroussi, A. Hasman, L. Pape-Haugaard, O. Saka, & S. K. Andersen (Eds.), *Studies in Health Technology and Informatics, Vol. 205: eHealth–For continuity of care* (pp. 905–909). Amsterdam: IOS Press.

Kuziemsky, C. E., & Peyton, L. A. (2016). Framework for understanding process interoperability and health information technology. *Health Policy and Technology, 5*(2), 196–203.

McDonald, C. J., Overhage, J. M., Tierney, W. M., Dexter, P. R., Martin, D. K., Suico, J. G., … Wodniak, C. (1999). The Regenstrief medical record system: A quarter century experience. *International Journal of Medical Informatics, 54*(3), 225–253.

McGlynn, E. A., Shekelle, P., & Hussey, P. (2008). *Developing, disseminating and assessing standards in the National Health Service*. Cambridge, UK: RAND Health.

McGowan, J. J., Cusack, C. M., & Poon, E. G. (2008). Formative evaluation: A critical component in EHR implementation. *Journal of American Medical Informatics Association, 15*(3), 297–301.

Mennemeyer, S. T., Menachemi, N., Rahurkar, S., & Ford, E. W. (2015). Impact of the HITECH act on physicians' adoption of electronic health records. *Journal of the American Medical Informatics Association, 23*(2), 375–379.

Niazkhani, Z., van der Sijs, H., Pirnejad, H., Redekop, W. K., & Aarts, J. (2009). Same system, different outcomes: Comparing the transitions from two paper-based systems to the same computerized physician order entry system. *International Journal of Medical Informatics, 78*(3), 170–181.

Novak, L., Brooks, J., Gadd, C., Anders, S., & Lorenzi, N. (2012). Mediating the intersections of organizational routines during the introduction of a health IT system. *European Journal of Information Systems, 21*(5), 552–569. doi: 10.1057/ejis.2012.2

Nykänen, P., Brender, J., Talmon, J., de Keizer, N., Rigby, M., Beuscart-Zephir, M. C., & Ammenwerth, E. (2011). Guideline for good evaluation practice in health informatics (GEP-HI). *International Journal of Medical Informatics, 80*(12), 815–827.

Pryor, T. A. (1988). The HELP medical record system. *MD Computing, 5*(5), 22–33.

Rozenblum, R., Jang, Y., Zimlichman, E., Salzberg, C., Tamblyn, M., Buckeridge, D., … Tamblyn, R. (2011). A qualitative study of Canada's experience with the implementation of electronic health information technology. *CMAJ : Canadian Medical Association Journal = Journal de l'Association médicale canadienne, 183*(5), E281–E288

Sheikh, A., Sood, H. S., & Bates, D. W. (2015). Leveraging health information technology to achieve the "triple aim" of healthcare reform. *Journal of the American Medical Informatics Association, 22*(4), 849–856.

Sherer, S. A., Meyerhoefer, C. D., Sheinberg, M., & Levick, D. (2015). Integrating commercial ambulatory electronic health records with hospital systems: An evolutionary process. *International Journal of Medical Informatics, 84*(9), 683–693.

Talmon, J., Ammenwerth, E., Brender, J., de Keizer, N., Nykänen, P., & Rigby, M. (2009). STARE-HI — Statement on reporting of evaluation studies in health informatics. *International Journal of Medical Informatics, 78*(1), 1—9.

Teich, J. M., Glaser, J. P., Beckley, R. F., Aranow, M., Bates, D. W., Kuperman, G. J., Ware, M. E., & Spurr, C. D. (1999). The Brigham integrated computing system (BICS): advanced clinical systems in an academic hospital environment. *International Journal of Medical Informatics, 54*(3), 197–208.

Vezyridis, P., & Timmons, S. (2014). National targets, process transformation and local consequences in an NHS emergency department (ED): A qualitative study. *BMC Emergency Medicine, 14*(1), 12. doi: 10.1186/1471-227X-14-12

Part III

Selected eHealth Evaluation Studies

Chapter 18

Value for Money in eHealth

Meta-synthesis of Current Evidence

Francis Lau

18.1 Introduction

Over the years a number of systematic reviews on studies that evaluated the economic return of eHealth investments have been published in the literature. Notable examples are the review on the financial effects of HIT by Low et al. (2013) based on 57 studies, the economics of HIT in medication management by O'Reilly, Tarride, Goeree, Lokker, and McKibbon (2012) based on 31 studies, and the scoping review of HIS on value for money by Bassi and Lau (2013) based on 42 studies. At a glance, these review findings seem favourable with over half of the studies showing positive economic returns. However, one should be mindful the studies were based on a diverse set of economic evaluation methods ranging from cost, outcome to full economic analysis done through modelling and field settings under different assumptions. Also the authors of these reviews have stressed the limitations of their findings. They include the heterogeneity of the eHealth systems examined, lack of detail on the system features, weak study designs with diverse costing/valuation methods and measures, and difficulty in generalizing the results. More importantly, not all of the studies were full economic evaluations and, thus, it was difficult to determine if the reported benefits were worth the investments. Few studies included the incremental cost of producing an extra unit of outcome and the long-term effect of the eHealth system.

In this chapter, three economic evaluation case studies that have been reported in the literature are presented to demonstrate value for money in eHealth. The three examples are: (a) a meta-synthesis of published eHealth economic reviews (section 18.2); (b) the cost-effectiveness and utility of computer-supported diabetes care in Ontario, Canada (section 18.3); and (c) the budget impact and

sustainability of system-wide human immunodeficiency virus (HIV) testing in the Veterans Affairs (VA) Administration in the United States (section 18.4). This is followed by a summary of the current state of evidence on eHealth economic evaluation for those involved in eHealth investment decisions (section 18.5).

18.2 Evidence on Value for Money in eHealth

This section examines the results of a meta-synthesis of the three published eHealth economic evaluation reviews by Low et al. (2013), Bassi and Lau (2013), and O'Reilly, Tarride, et al. (2012). The intention was to combine these reviews to make sense of the current state of evidence on value for money in eHealth investments. To do so, first the three original reviews were reanalyzed to reconcile the mixed findings. Then the focus turned to the full economic evaluation studies from these reviews that were published between 2000 and 2010 in order to gain insights on the economic return for specific types of eHealth systems.

18.2.1 Synopsis of Economic Review Findings

The review by Low and colleagues (2013) found that 75.4% (or 43 out of 57) of their studies had reported financial benefits in the form of revenue gains and cost savings to stakeholders. The eHealth systems in question were: 42.1% (24/57) CPOE/CDS (computerized provider order entry/clinical decision support); 45.6% (26/57) EHR; 8.8% (5/57) HIE; and 3.5% (2/57) combined. The proportions of systems with reported benefits included: 88.4% (14/17) outpatient EHR; 69.2% (9/13) outpatient CPOE/CDS; 60.0% (6/10) inpatient CPOE/CDS; and 75.0% (3/4) Emergency Department HIE.

The review by Bassi and Lau (2013) found that 69.7% (or 23 out of 33) of their high-quality studies (quality score ≥ 8/10) had reported positive returns. The eHealth systems in question were: 21.2% (7/33) primary care EMR; 18.2% (6/33) CPOE; 15.2% (5/33) medication management; 15.2% (5/33) immunization; 12.1% (4/33) HIS; 9.1% (3/33) disease management; 6.1% (2/33) clinical documentation; and 3.0% (1/33) HIE. The proportions of systems with positive returns included: 71.4% (5/7) primary care EMR; 50.0% (3/6) CPOE; 100% (5/5) medication management; 60.0% (3/5) immunization; 75.0% (3/4) HIS; 100% (3/3) disease management; and 100% (1/1) HIE. The remaining two clinical documentation systems had inconclusive results.

The review by O'Reilly, Tarride, et al. (2012) had 31 studies but only narrative descriptions were reported because of the heterogeneity of the settings, systems and methods involved. While the review was on medication management, the HITS evaluated varied and were mostly CPOE, CDS, MAR (medication administration record), and combined systems (67.7% or 21/31 studies) with the remaining as barcode, EMR or cardiopulmonary resuscitation (CPR), surveillance and ePrescribing systems. The authors did not summarize the proportion of studies with economic benefits but a tabulation from the narrative tables in the review showed that 67.7% or 21 out of 31 studies had reported some cost benefits.

18.2.2 Meta-synthesis of Full Economic Evaluation Studies

Combined, the three reviews had a total of 121 evaluation studies published during the period between 1993 and 2010. To make sense of the review findings, a reanalysis of all of the studies was conducted by reconciling for duplicates, selecting only those published in English between 2000 and 2010, then grouping them by economic analysis method. This reanalysis led to a combined list of 81 unique studies, of which only 19 or 23.5% were considered full economic evaluation. These 19 studies were then synthesized to provide an economic evidence base for eHealth systems in primary care EMR, medication management, CPOE/CDS, institutional HIS, disease management, immunization, documentation and HIE, as defined by Bassi and Lau (2013).

A summary of the 19 studies by eHealth system, author-year, time frame, options, cost, outcome, comparison method, results and interpretation is shown in the Appendix. Of these 19 studies, seven were on primary care EMRs, three on medication management, three on CPOE/CDS, two on institutional HIS, and one each on disease management, immunization, documentation and HIE, respectively. For designs, 78.9% (15/19) were field studies and 21.1% (4/19) were simulations. For methods, 73.7% (14/19) of the studies were cost-benefit, 21.1% (4/19) cost-effectiveness, and 5.2% (1/19) cost-consequence analysis. Two studies also included cost-minimization as a second method. For valuation, 52.6% (10/19) of the studies included some type of discounting and/or inflation to determine the present dollar value. Of the 19 studies, only 36.9% (7/19) included one- or multi-way sensitivity analysis, 21.1% (4/19) reported the incremental cost-effectiveness ratio (ICER), and 5.3% (1/19) included quality-adjusted life years (QALY).

For results, there were positive returns on investment in 100% (7/7) of the primary EMR, 66.7% (2/3) of CPOE/CDS, 100% (2/2) of institutional EHR, and 100% for each of the disease management (1/1), immunization (1/1) and HIE (1/1) systems. The three medication management systems and the one documentation system did show positive returns but only when specific conditions were met. A closer examination of the results revealed that the positive returns from the primary care EMR studies were mainly productivity-related in terms of cost savings and increased revenues, with little mention of tangible improvement in health outcomes. One CPOE study had mixed results in that the simulated operating margins from CPOE adoption were positive over time for large urban hospitals but for not rural or critical access hospitals. The three medication management studies were inconclusive as they were dependent on certain contextual factors. For instance, Wu, Laporte, and Ungar (2007) showed an incremental cost-effectiveness ratio (ICER) of $12,700 per adverse drug event (ADE) averted. That translated to 32.3 ADEs averted per year or 261 events averted over 10 years, but the estimates depended on the base rate of adverse drug events, system and physician costs and the ability to reduce ADEs. Fretheim, Aaserud, and Oxman (2006) found the cost of thiazide intervention to be twice the cost savings in year-1 before modest savings could be projected in year-2 by expand-

ing the intervention into a national program. Similarly, the clinical documentation study by Kopach, Geiger, and Ungar (2005) showed ICER of $0.331 per day in average discharge note completion time, depending upon physician utilization volume and length of the study.

Based on the results of this small set of full economic analysis studies there is some evidence to suggest value for money in eHealth investments in selected healthcare domains and types of systems. However, the number of studies is small and caution is needed when generalizing these results to other settings.

18.3 Computer-supported Diabetes Care in Ontario

This section presents a set of economic evaluation modelling and field studies on diabetes care done in the Canadian Province of Ontario over the past 15+ years that began around the year 1999. These include: (a) the application of the Ontario Diabetes Economic Model (ODEM) in the COMPETE-II randomized trial; and (b) a mega-analysis on optimizing chronic disease management that includes electronic tools for diabetes care. These studies are described below.

18.3.1 Application of ODEM in COMPETE-II Trial

The ODEM is a simulation model that uses a set of parametric risk equations, based on specific patient characteristics, to predict the cost and occurrence of diabetes-related complications, life expectancy and quality-adjusted life years (QALYS) over a 40-year time horizon. The ODEM is an adaptation of the United Kingdom Prospective Diabetes Study (UKPDS) Outcomes Model, which was developed with data from the UKPDS conducted as a randomized trial in the 1970s (Clarke et al., 2004).

Holbrook and colleagues (2009) conducted the COMPETE-II study in Ontario as a pragmatic randomized trial during 2002 and 2003. Its objective was to determine if electronic decision support and shared information with diabetic patients could improve their care in the community setting. The study was conducted in three Ontario regions with 46 primary care practices and adult patients under their care. The study results were then applied as inputs to the ODEM in a modelling study to estimate the long-term quality of life and cost implications (O'Reilly, Holbrook, Blackhouse, Troyan, & Goeree, 2012). The key aspects of the two studies are summarized below in terms of the diabetes cohort, intervention, economic analysis and projected benefit.

> *Diabetes Cohort* – The study had 511 adult type-2 diabetic patients, with 252 randomized to the intervention group and 258 to the control group. The mean follow-up time was 5.9 months and the median time since diagnosis of diabetes was 5.9 years. Key risk factors from the trial were used as input to the ODEM such as HbA1c (glycated hemoglobin test), systolic blood pressure, cholesterol and smoking status. The costs of resource use and diabetes-related

complications in the ODEM were derived from a prospective cohort of 734,113 diabetic patients over a 10-year period representing 4.4 million patient-years in Ontario.

Intervention – An individualized electronic decision support (DS) and reminder system for diabetes care was implemented in three Ontario regions for use by 46 primary care practices over a one-year period. The intervention included a Web-based diabetes tracker template for shared access by providers and patients, an automated phone reminder for patients, and a colour tracker page mailed to patients. The diabetes tracker template was interfaced with the EMR and allowed the display and monitoring of 13 risk factors, specifically blood pressure, cholesterol, HbA1c, foot exam, kidney, weight, physical activity, smoking, eye exam, acetylsalicylic acid or equivalent, ACE inhibitors, and flu shot. The automated phone reminder system prompted patients every month to follow up on medications, labs and physician visits. The colour tracker page was mailed to patients four times a year and was to be taken to physician appointments.

Economic Analysis – The long-term cost-effectiveness of the shared DS and reminder system was examined. The respective economic evaluation components are summarized below.

- Perspective – Ontario Ministry of Health;
- Options – A shared DS and reminder system versus usual care;
- Time Frame – 40-year time horizon after the 12-month study in 2002-03, assuming a one-year treatment effect at 5% discount rate in 2010 Canadian dollars;
- Input Costs – Program implementation costs and projected diabetes complications. Program costs included tracker development and testing, ongoing project management, and required IT infrastructure;
- Outcomes – Intermediate outcomes (HbA1c, blood pressure, cholesterol and smoking), life years, quality-adjusted life years (QALYS), incremental costs, and ICER;
- Comparison of Options – Cost-effectiveness analysis to compare lifetime effects of DS and reminder system versus usual care in expected costs per patient, life years, QALYS and ICER. Sensitivity analysis to compare lifetime effects of program and treatment effect duration of one, five and 10 years, and discount rates of 0%, 3% and 5%.

Projected Benefit – The intervention reduced HbA1c by 0.2 and systolic blood pressure by 3.95 mmHG, and an overall relative risk

reduction of 14% in the need for amputation. The total cost of the intervention was $483,699, at a mean lifetime cost of $1,912 per patient receiving the intervention. The ODEM estimated the disease management costs to be $61,340 and $61,367 for the intervention and control groups, respectively, at an incremental cost of –$26 per patient. The avoidance of complications would gain an additional 0.0117 QALYs, and an estimated ICER of $156,970 per life year and $160,845 per QALY. Sensitivity analysis showed an increase of 260% in QALYs from 0.0117 to 0.0421 when patients were treated for five years due to reduced downstream complications, at an ICER of $186,728. When patients were treated for 10 years there was a sixfold increase in QALYs gained, at an ICER of $173,654. Overall, the intervention led to slight improvement in short-term risk factors and moderate improvement in long-term health outcomes. To do so, the intervention had to be highly efficient and effective in its costs and care processes.

18.3.2 Optimizing Chronic Disease Management Mega-analysis

In 2013 Health Quality Ontario (HQO) published a mega-analysis series drawn from 15 reports on the economic aspects of community-based chronic disease management (CDM) interventions (HQO, 2013). The chronic diseases examined were diabetes, chronic obstructive pulmonary disease (COPD), coronary artery disease, and congestive heart failure. The CDM interventions included discharge planning, continuity of care, in-home care, specialized nursing practice, and electronic tools (eTools) for health information exchange (HIE). The eTools for HIE component of this mega-analysis in diabetes care is summarized below in terms of the diabetes cohort, intervention, economic analysis and projected benefit.

Diabetes Cohort – Adult patients with type-2 diabetes-related physician visits or one hospital admission within two years between 2006 and 2011 were included as the Ontario cohort. For each patient, their resource use and mean 90-day total costs by sector were estimated from the Ontario administrative databases. These included emergency department visits, acute inpatient and same-day surgery costs, other hospital costs, long-term care, home care and physician visits, lab costs and drug costs. The EQ-5D (European Quality of Life 5 Dimensions) values were used as the utility estimates for changes in quality of life from hospitalizations during the study period. The mean EQ-5D value of 0.77 derived from 3,192 patients in the UKPDS (Clarke, Gray, & Holman, 2002) was used as the baseline utility estimate for the Ontario cohort. The mean EQ-5D value of 0.54 was used as a proxy measure for hospitalization, based on the study on severe hypoglycemia in diabetics by Davis et al. (2005). Patients in the Ontario cohort who

were hospitalized were assigned the utility value of 0.54 over their average length of stay. For the intervention group, a 0.85 relative difference in hospitalization from an eTools for HIE field trial by Kahn, MacLean, and Littenberg (2010) was applied as a result of improved quality of life, thereby reducing the proportion of patients hospitalized.

Intervention – The Vermont Diabetes Information System (VDIS) developed by MacLean, Littenberg, and Gagnon (2006) was used as the model eTool for HIE intervention. The VDIS is a decision support system that sends lab results, reminders and alerts to primary care providers and their patients with diabetes. Quarterly population reports were also available to providers for peer comparison. A randomized trial by MacLean, Gagnon, Callas, and Littenberg (2009) showed that VDIS improved lab monitoring of diabetic patients in primary care but not physiologic control. For cost, the VDIS vendor quoted a one-time software cost of $5,000 and an annual maintenance cost of $2,500 per laboratory. The annual cost to receive VDIS information was $6,000 per physician and $48 per patient in 2012 Canadian dollars. The per-patient costs were dependent on physician roster size and disease prevalence. Since no eTools for HIE were in regular use in Ontario at the time of the mega-analysis, the proportion of diabetic patients that could benefit from HIE was assumed to be 100%.

Economic Analysis – The projected cost-effectiveness of the modelled eTools for HIE in community-based care were examined. The respective economic evaluation components are summarized below.

- Perspective – The Ontario provincial health ministry level (i.e., Ministry of Health and Long-Term Care);
- Options – Hypothetical adoption of VDIS as the eTools for HIE versus usual care with no HIE;
- Time Frame – A five-year horizon with an annual 5% discount rate inflated to 2012 Canadian dollars; duration of benefit assumed to be 32 months based on the literature;
- Input Costs – Estimated resource use costs with or without hospitalization for the Ontario cohort based on administrative data over a five-year period. Estimated one-time intervention costs covered and ongoing VDIS costs for 211 labs, 11,902 physicians and 85 diabetic patients per physician;
- Outcomes – Proportion of hospitalized patients based on severe hypoglycemia as a proxy measure from the literature and QALYS

with or without hospitalization based on EQ-5D values as utility estimates from the literature;
- Comparison of Options – Cost-effectiveness analysis to compare eTools with usual care options in cost per patient, QALYs per patient, and ICER. Sensitivity analysis to compare changes in relative difference of hospitalization and emergency department visits, and marginal ongoing costs in the intervention group.

Projected Benefit – The cost-effectiveness analysis showed that the cost per patient was $29,889 with eTools versus $30,226 with usual care. The QALYs per patient was 2.795 with eTools versus 2.789 with usual care. The ICER was –$337 per patient. The sensitivity analysis showed the model was sensitive to changes in resource use and intervention cost. For instance, a relative difference of 0.75 in hospitalization for the intervention would change the ICER to –$1,228, where a relative difference of 0.95 would change the ICER to $654. A marginal cost of $74 in ongoing cost for the intervention would change the ICER to –$724, but a marginal cost of $233 would change the ICER to $639. Overall, the intervention was found to be less costly and more effective when compared with usual care.

18.4 System-wide HIV Testing in Veterans Affairs (VA) Administration

In 1998 the United States VA Administration launched the Quality Enhancement Research Initiative (QUERI) to improve the performance of the VA healthcare system and the consequent quality of care for its veterans (Smith & Barnett, 2008). In that initiative, QUERI researchers collaborated with VA leaders and staff to implement evidence-based practice as the routine standard of care through a six-step process:

1 Identify high-risk/volume diseases or problems.

2 Identify best practices.

3 Identify deviations from current practices and outcomes.

4 Identify and implement interventions to promote best practices.

5 Document that best practices improved outcomes.

6 Document that outcomes were associated with improved health-related quality of life.

For step-4 above, the implementation efforts followed a sequence of phases from a single-site pilot project to a small-scale multisite trial, followed by a large-scale multi-region trial to a final system-wide rollout. An integral part of the initiative was the use of policy cost-effectiveness and budget impact analysis in single-site and multisite trials to determine the economic return. This section describes a case study on HIV testing at the VA Administration in terms of the multi-component intervention program, different implementation phases it went through over the years, and budget impact analysis done on the program.

18.4.1 Multi-component Intervention Program

The multi-component intervention was made up of computerized decision support, audit-feedback, provider activation and organizational level change (Goetz et al., 2008). The computerized decision support was a real-time clinical reminder that identified patients at increased risk for HIV infections and prompted healthcare providers to offer HIV testing to these patients. The clinical reminder was triggered by the presence of a set of predefined criteria such as prior Hepatitis B or C infection, sexually transmitted disease, drug use, homelessness and specific behavioural risk factors (e.g., excessive alcohol use, multiple sexual partners, body piercing). These data elements were automatically extracted from the VA EMR during the patient visit. Once triggered, the provider had to address the reminder by ordering an HIV test, asking the test to be done elsewhere, recording that the patient was either not competent to consent to testing or had refused testing.

An audit-feedback system was developed to inform providers of their performance in HIV evaluation and testing rates of at-risk patients at the clinic level. The reports were distributed to clinical leaders and clinic managers via e-mail on a quarterly basis. Provider activation included the use of academic detailing, social marketing and educational materials to engage both providers and patients in the initiative. Academic detailing involved one-on-one sessions in person and ad-hoc site visits with project staff to discuss the need for and benefits of HIV testing. Social marketing involved the recruitment of physician and nurse leaders to encourage HIV testing at the clinic. Educational materials included information handouts, pocket cards and posters to inform providers and patients on the need and criteria for, and process and implications of HIV testing. Change at the organizational level involved the removal of barriers to HIV testing, such as the inclusion of streamlined pretest counselling that only took two to three minutes, and post-test phone notification and brief counselling of negative test results.

18.4.2 Program Implementation and Evaluation

Goetz and colleagues (2008) conducted a pre-post intervention study from 2004 to 2006 to determine if the multi-component intervention program would increase the rate of HIV testing. Five VA facilities took part in the study with two

receiving the intervention and three as controls. The HIV testing rate and the number of newly diagnosed cases in the year before and after implementing the intervention were compared. Patient, provider and facility-level factors that could influence testing performance were also examined. These included patient subgroups with different demographics and risk factors, proportions of at-risk patients tested by primary providers, as well as the prevalence of at-risk patients and annual patient load at the facility. The results showed 36,790 untested patients with HIV risk factors from the intervention sites and 44,577 patients from the control sites were considered in the study. The adjusted rate of HIV testing at the two intervention sites increased from 4.8% to 10.8% and from 5.5% to 12.8%, and the number of newly diagnosed cases increased from 15 to 30 after implementing the intervention. There was no change in the control sites during the same period. Overall the intervention was considered effective in increasing the HIV testing rate and the detection of new cases.

Sustainability of the Intervention – Goetz et al. (2009) evaluated the sustainability of increased HIV testing after implementing the multi-component intervention program in 2005. The intervention was implemented in month-1 of the intervention year 2005 then continued for the subsequent 11 months. During the intervention year the study team supported the provider activation component of the intervention that included academic detailing, social marketing, and provision of educational materials. In year-2, or the sustainability year, the responsibility for provider activation was transferred to the clinic. During this period the clinical reminders continued to be used, the quarterly feedback reports were managed by clinical leaders, and provider education activities were reduced and merged with regular staff meetings. Further organizational changes broadened the number of providers who could order the test, eased the documentation requirements and continued with the pretest and post-test counselling. The results showed the monthly adjusted testing rate increased from 2% at baseline to 6% by the end of the intervention year. Then the rate declined to 4% by the end of the sustainability year. The testing rate for persons newly exposed to the intervention increased during the intervention and sustainability years. The attenuation effect in the sustainability year was caused by the increase in the proportion of visits by untested patients despite prior exposures to the intervention. The percentage of patients who received HIV testing was 5.0% in the pre-intervention year, 11.1% in the intervention year, and 11.6% in the sustainability year. Overall, the intervention was considered sustainable, especially in patients during their early contacts with the healthcare system.

Scalability of the Intervention – Goetz and colleagues (2013) also evaluated the scalability of the multi-component intervention in routine HIV testing and the level of support needed. A one-year three-arm quasi-experimental study was conducted with central support, local support, and no support (i.e., control) provided to different VA primary care sites in three geographic regions. All sites had access to the real-time clinical reminder system. With central support, the study team provided quarterly audit-feedback reports, provider activation and ongoing support including site visits. With local support, the sites had only a single conference call 30 days after the initial site visit. The control sites had no contact with the study team. The clinical reminder was initially risk-based for all sites in the first six months of the study, then became routine for all patients in the following six months. In phase-1, the adjusted rate of risk-based testing increased by 10.1%, 5.6% and 0.4% in the central, local and control sites, respectively. In phase-2, the adjusted rate of routine testing increased by 9.2%, 6.3% and 1.1% in the central, local and control sites. By the end of the study, 70% to 80% of VA patients had been offered an HIV test. Overall, the multi-component intervention program was considered scalable in reaching the goal of all VA patients being aware of their HIV status as part of routine clinical visits.

18.4.3 Budget Impact Analysis

Anaya, Chan, Karmarkar, Asch, and Goetz (2012) conducted a budget impact study to examine the facility-specific costs of HIV testing and care for newly identified HIV patients. The study was based on the multi-component HIV intervention program discussed above, that was implemented as a pre-post quasi-experimental trial in five Veterans Health Administration facilities (Goetz et al., 2008). A budget impact model was developed to estimate the costs of HIV testing that included the costs of pretest counselling, HIV testing rates, and treatment of identified HIV patients. The budget impact model, intervention, economic analysis and projected benefits are summarized below.

Budget Impact Model – The model was developed to estimate the costs of HIV testing in a single VA facility in the primary care setting. Two HIV providers were consulted to establish relevant model end points. They covered physician and nurse staffing costs, laboratory costs, and the costs of antiretroviral therapy (ART) for different levels of HIV disease progression based on Cluster of Differentiation 4 (CD4) counts. The model included quarter-to-quarter changes in patient status, loss to follow-up and deaths that occurred in a period. It covered the costs of tested and untested patients of known and unknown HIV status who received care in

a single facility over eight three-month periods. A hypothetical cohort of 20,000 adult patients was used, with a prevalence of 9.2% as having already been tested, 200 as known HIV patients under care, three minutes of extra nursing time, and a 2.1% annual baseline HIV testing rate in untested patients.

Intervention – The multi-component intervention program consisted of a real-time electronic clinical reminder for HIV testing, audit-feedback reports, provider activation and patient-provider education.

Economic Analysis – The budget impact of expanded HIV testing in a primary care setting were examined. The respective economic evaluation components are summarized below.

- Perspective – The integrated VA healthcare system that offer both HIV testing and care;
- Options – Expanded HIV testing rate of 15% versus baseline rate of 2.1%;
- Time Frame – A two-year horizon in eight three-month quarterly periods;
- Input Costs – Personnel and laboratory costs, and ART costs from different levels of HIV disease progression based on CD4 count, tracked on a quarterly basis;
- Outcomes – HIV testing rates, number and percent of HIV-positive patients at different CD4 levels;
- Comparison of Options – Budget impact on expanded HIV testing from 2.1% to 15% at 0.45% positive test rate; sensitivity analysis with HIV testing rates from 15% to 30%, positive test rate from 0.45% to 1%, and pretest nursing time activities from three to five minutes.

Projected Benefit – The expansion of HIV testing from 2.1% to 15% annually led to the identification of 21 additional HIV-positive patients over two years at a cost of $290,000. Over 60% of this cost was to provide ART to newly diagnosed patients. Quarterly ART costs increased from $10,000 to more than $60,000 over two years with more HIV patients identified and treated with ART. In sensitivity analysis, serodiagnostic and annual HIV testing rates had the greatest cost impact. Overall, expanded HIV testing led to increased initial costs, mostly due to ART treatment for new patients. Using a $50,000 per QALY threshold, expanded HIV testing was cost-effective based on a total cost of $80,000 over two years for testing, and $290,000 for testing and care for 21 additional HIV patients.

18.5 Summary of Economic Evidence in eHealth

Overall, our meta-synthesis of the three published eHealth economic evaluation reviews showed that there is value for money in eHealth investment. However, the evidence varied depending on the domains, contexts and systems involved. This evidence is strong in primary care EMR as all seven full economic analysis studies had positive returns. For CPOE/CDS, institutional EHR, disease management, immunization and HIE systems, while there is evidence of positive returns it is much weaker since they are only based on a small number of modelling and field studies. For medication management and documentation systems, the evidence is weak to inconclusive since the positive return is contingent on the interplay of different socio-organizational, technical and external factors.

The development and validation of the ODEM and its application in the COMPETE-II and HIE studies in Ontario, Canada showed that computer-supported diabetes care could be cost-effective but required a great deal of effort to implement and maintain the interventions. With the electronic diabetes tracker, there was a modest benefit in achieving process outcome targets in the short term, and some gain in QALYS with reduced complications in the long term. However, the projected economic return was contingent on the precision of the ODEM parameter estimates such as the disease prevalence, resource use and costs, complication rates, and provider EMR adoption behaviours. The HIE modelling study was cost-effective in sharing patient information but it assumed 100% adoption of the eTools by all primary care providers in the province. Similarly, the multi-component HIV testing program in Veterans Affairs Administration in the United States showed that computerized HIV testing was cost-effective when combined with patient-provider activation and organizational policies. Once implemented, the risk-based testing program was shown to be sustainable with more streamlined support and eventually scalable as a routine practice in the organization. The ICER and gain in QALYS were considered good return on value.

References

Anaya, H. D., Chan, K., Kamarkar, U., Asch, S. M., & Goetz, M. B. (2012). Budget impact analysis of HIV testing in the VA healthcare system. *Value in Health, 15*(8), 1022–1028.

Bassi, J., & Lau, F. (2013). Measuring value for money: A scoping review on economic evaluation of health information systems. *Journal of American Medical Informatics Association, 20*(4), 792–801.

Clarke, P., Gray, A., & Holman, R. (2002). Estimating utility values for health states of type 2 diabetic patients using the EQ-5D (UKPDS 62). *Medical Decision Making, 22*(4), 340–349.

Clarke, P. M., Gray, A. M., Briggs, A., Farmer, A. J., Fenn, P., Stevens, R. J., … UK Prospective Diabetes Study (UKDPS) Group. (2004). A model to estimate the lifetime health outcomes of patients with type 2 diabetes: the United Kingdom Prospective Diabetes Study (UKPDS) outcomes model (UKPDS no. 68). *Diabetologia, 47*(10), 1747–1759.

Davis, R. E., Morrissey, M., Peters, J. R., Wittrup-Jensen, K., Kennedy-Martin, T., & Currie, C. J. (2005). Impact of hypoglycaemia on quality of life and productivity in type 1 and type 2 diabetes. *Current Medical Research and Opinion, 21*(9), 1477–1483.

Fretheim, A., Aaserud, M., & Oxman, A. D. (2006). Rational prescribing in primary care (RaPP): economic evaluation of an intervention to improve professional practice. *PLoS Medicine, 3*(6), e216.

Goetz, M. B., Hoang, T., Bowman, C., Knapp, H., Rossman, B., Smith, R., … Asch, S. M. (2008). A system-wide intervention to improve HIV testing in the Veterans health administration. *Journal of General Internal Medicine, 23*(8), 1200–1207.

Goetz, M. B., Hoang, T., Henry, R., Knapp, H., Anaya, H. D., Gifford, A. L., & Asch, S. M. (2009). Evaluation of the sustainability of an intervention to increase HIV testing. *Journal of General Internal Medicine, 24*(12), 1275–1280.

Goetz, M. B., Hoang, T., Knapp, H., Burgess, J., Fletcher, M. D., Gifford, A. L., & Asch, S. M. (2013). Central implementation strategies outperform local ones in improving HIV testing in Veterans healthcare administration facilities. *Journal of General Internal Medicine, 28*(10), 1311–1317.

Health Quality Ontario. (2013). Optimizing chronic disease management mega-analysis: Economic evaluation. *Ontario Health Technology Assessment Series, 13*(13), 1–148.

Holbrook, A., Thabane, L., Keshavjee, K., Dolovich, L., Bernstein, B., Chan, D., … Gerstein, H. (2009). Individualized electronic decision support and reminders to improve diabetes care in the community: COMPETE II randomized trial. *Canadian Medical Association Journal, 181*(102), 37–44.

Jones, S. S., Rudin, R. S., Perry, T., & Shekelle, P. G. (2014). Health information technology: An updated systematic review with a focus on meaningful use. *Annals of Internal Medicine, 160*(1), 48–54.

Khan, S., MacLean, C. D., & Littenberg, B. (2010). The effect of the Vermont diabetes information system on inpatient and emergency department use: Results from a randomized trial. *Health Outcomes in Research in Medicine, 1*(1), e61–e66.

Kopach, R., Geiger, G., & Ungar, W. J. (2005). Cost-effectiveness analysis of medical documentation alternatives. *International Journal of Technology Assessment in Health Care, 21*(1), 126–131.

Low, A. F. H., Phillips, A. B., Ancker, J. S., Patel, A. R., Kern, L. M., & Kaushal, R. (2013). Financial effects of health information technology: a systematic review. *American Journal of Managed Care, 19*(10 Spec No), SP369–SP376.

MacLean, C. D., Littenberg, B., & Gagnon, M. (2006). Diabetes decision support: Initial experience with the Vermont diabetes information system. *American Journal of Public Health, 96*(4), 593–595.

MacLean, C. D., Gagnon, M., Callas, P., & Littenberg, B. (2009). The Vermont diabetes information system: A cluster randomized trial of a population-based decision support system. *Journal of General Internal Medicine, 24*(12), 1303–1310.

O'Reilly, D., Holbrook, A., Blackhouse, G., Troyan, S., & Goeree, R. (2012). Cost-effectiveness of a shared computerized decision support system for diabetes linked to electronic medical records. *Journal of American Medical Informatics Association, 19*(3), 341–345.

O'Reilly, D., Tarride, J. E., Goeree, R., Lokker, C., & McKibbon, K. A. (2012). The economics of health information technology in medication management: a systematic review of economic evaluations. *Journal of American Medical Informatics Association, 19*(3), 423–438.

Smith, M. W., & Barnett, P. G. (2008). The role of economics in the QUERI program: QUERI series. *Implementation Science, 3*(20). doi: 10.1186/1748-5908-3-20

Wu, R. C., Laporte, A., & Ungar, W. J. (2007). Cost effectiveness of an electronic medication ordering and administration system in reducing adverse drug events. *Journal of Evaluation in Clinical Practice, 13*(3), 440–448.

Appendix

Summary of Economic Evaluation Findings

Study	Time Frame	Options	Cost	Outcome	Comparison	Results	Interpretation
Primary Care EMR							
Block (2008) USA	3 years 2005-2007	EMR system vs. paper for net cost and benefit	EMR start-up costs in year-1 with hardware, software, training, implementation, data migration, training, support, lab interface, appointment reminder system in year-2	Reduced staffing cost, increased productivity and billing	Comparison of cost and benefit and net return on investment over 3 years	Year-1 cost saving of $5,500 per month in payroll and benefits, in year-2 increased savings to $6,800 per month; net annual returns of $6,200, $59,250 and $96,150 in years 1, 2 and 3	+ Positive net return on investment; minor ongoing IT costs not included (<$500/physician per year)
Grieger et al. (2007) USA	2.5 years 2003-2005	EMR system vs. paper for net cost and benefit	Year-1 capital cost for hardware, software, technical support, training; ongoing operating expenses	Reduced times for chart pull, new charts, filing, transcriptions; reduced support staff, patient cycle time, days in accounts receivable; improved billing	Comparison of cost and benefit and net return on investment	Year-1 expenses were $509,539, ongoing annual cost year-2 were $114,016; initial costs recovered in 16 months with ongoing savings of $279,524 or $9,983 per provider	+ Positive net return on investment; neutral impact on efficiency and billing
Kumar and Bauer (2011) USA	Hypothetical 5 year time period	EMR system for costs and benefits with no options	Reported software, hardware, installation, training, ongoing maintenance, support staffing, and productivity loss from literature (Wang et al., 2003; Miller et al., 2005); 10% discount rate	Reported savings from transcription, chart pull, malpractice insurance; and chart capture from literature	Simulation with Simulation with random inputs for minimum, maximum and mode cost-benefit estimates; net benefit and present value scenarios at 10%, 15% and 25% discount rates	Net present values of $124,725, $106,635, $79,395 at 10%, 15%, 25% discount over 5 years; worst case scenario with maximum costs, lowest benefits at 25% had (-$9,462)	+ Positive net present value demonstrated across different assumptions, but sensitive to local organizational factors

Summary of Economic Evaluation Findings

Study	Time Frame	Options	Cost	Outcome	Comparison	Results	Interpretation
Primary Care EMR							
Miller et al. (2005) USA	1 year 2004-2005	EMR system for costs and benefits, with no options	Estimated one time and ongoing EMR, productivity loss, staffing costs from interviews, observations, reports and contracts	Estimated efficiency savings and gains, and revenue enhancement from reports, observations and interviews; averaged per FTE provider	Comparison of cost and benefit per FTE provider, and payback period	Average payback in 2.5 years, $23,000 net benefits per FTE provider per year, some could not recover costs quickly, faced financial risks	+ Need incentives and support services to improve quality of care
Patil et al. (2008) USA	8 years with 4 years each for pre-post periods 1998-2005	Manual vs. EMR system for cost and productivity	Historical costs for manual transcription and EMR implementation and maintenance, with 3% allocation for annual EMR cost, adjusted for inflation in 2006 US dollars	Average net revenue per encounter and per provider over 4 years in pre-post EMR, based on total revenue, transcription cost, EMR cost, encounter volume, and number of providers, adjusted for inflation to 2006 US dollars, extrapolated to 4 years post-EMR with average savings	Comparison of cost/revenue per encounter and per provider, with and without sunk and residual transcribing costs	Average cost saving $3.09 per encounter, increased revenue $117.88 per encounter and $184,627 per provider, start-up EMR cost $10,329 per provider	+ less cost savings if sunk and residual transcribing costs G9
Simon and Simon (2006) USA	1 year not stated, assumed 2005	EMR vs. paper for net benefit/cost	Hardware, software, implementation, training, including discounts and incentives	Tangible cost savings in reduced staffing for managing paper chart, transcription and improved claims	Comparison of costs and benefits, return on investment	Costs of $213,083 and benefits of $657,500, return on investment of 308%	+ Excluded lost productivity from implementation and intangible benefits

Summary of Economic Evaluation Findings

Study	Time Frame	Options	Cost	Outcome	Comparison	Results	Interpretation
Primary Care EMR							
Wang et al. (2003) USA	5 years not stated, assumed 2002-2006	EMR vs. paper for net benefit/cost	System and induced costs from historical records, experts and literature, discounted at 5% in 2002 US dollars	Estimated cost savings, increased revenues (payer independent, capitation, fee for service), discounted at 5% in 2002 US dollars	Net benefit per provider, with 1, 2 and 5-way sensitivity analysis	Net benefit $86,400 per provider over 5 years	+ based on proportion of patient capitation
Block (2008) USA	3 years 2005-2007	EMR system vs. paper for net cost and benefit	EMR start-up costs in year-1 with hardware, software, training, implementation, data migration, training, support, lab interface, appointment reminder system in year-2	Reduced staffing cost, increased productivity and billing	Comparison of cost and benefit and net return on investment over 3 years	Year-1 cost saving of $5,500 per month in payroll and benefits, in year-2 increased savings to $6,800 per month; net annual returns of $6,200, $59,250 and $96,150 in years 1, 2 and 3	+ Positive net return on investment; minor ongoing IT costs not included (<$500/physician per year)
Grieger et al. (2007) USA	2.5 years 2003-2005	EMR system vs. paper for net cost and benefit	Year-1 capital cost for hardware, software, technical support, training; ongoing operating expenses	Reduced times for chart pull, new charts, filing, transcriptions; reduced support staff, patient cycle time, days in accounts receivable; improved billing	Comparison of cost and benefit and net return on investment	Year-1 expenses were $509,539, ongoing annual cost year-2 were $114,016; initial costs recovered in 16 months with ongoing savings of $279,524 or $9,983 per provider	+ Positive net return on investment; neutral impact on efficiency and billing

Summary of Economic Evaluation Findings

Study	Time Frame	Options	Cost	Outcome	Comparison	Results	Interpretation
Primary Care EMR							
Kumar and Bauer (2011) USA	Hypothetical 5 year time period	EMR system for costs and benefits with no options	Reported software, hardware, installation, training, ongoing maintenance, support staffing, and productivity loss from literature (Wang et al., 2003; Miller et al., 2005); 10% discount rate	Reported savings from transcription, chart pull, malpractice insurance; and chart capture from literature	Simulation with Simulation with random inputs for minimum, maximum and mode cost-benefit estimates; net benefit and present value scenarios at 10%, 15% and 25% discount rates	Net present values of $124,725, $106,635, $79,395 at 10%, 15%, 25% discount over 5 years; worst case scenario with maximum costs, lowest benefits at 25% had (-$9,462)	+ Positive net present value demonstrated across different assumptions, but sensitive to local organizational factors
Miller et al. (2005) USA	1 year 2004-2005	EMR system for costs and benefits, with no options	Estimated one time and ongoing EMR, productivity loss, staffing costs from interviews, observations, reports and contracts	Estimated efficiency savings and gains, and revenue enhancement from reports, observations and interviews; averaged per FTE provider	Comparison of cost and benefit per FTE provider, and payback period	Average payback in 2.5 years, $23,000 net benefits per FTE provider per year, some could not recover costs quickly, faced financial risks	+ Need incentives and support services to improve quality of care

Summary of Economic Evaluation Findings

Study	Time Frame	Options	Cost	Outcome	Comparison	Results	Interpretation
Primary Care EMR							
Patil et al. (2008) USA	8 years with 4 years each for pre-post periods 1998-2005	Manual vs. EMR system for cost and productivity	Historical costs for manual transcription and EMR implementation and maintenance, with 3% allocation for annual EMR cost, adjusted for inflation in 2006 US dollars	Average net revenue per encounter and per provider over 4 years in pre-post EMR, based on total revenue, transcription cost, EMR cost, encounter volume, and number of providers, adjusted for inflation to 2006 US dollars, extrapolated to 4 years post-EMR with average savings	Comparison of cost/revenue per encounter and per provider, with and without sunk and residual transcribing costs	Average cost saving $3.09 per encounter, increased revenue $117.88 per encounter and $184,627 per provider, start-up EMR cost $10,329 per provider	+ less cost savings if sunk and residual transcribing costs G9
Simon and Simon (2006) USA	1 year not stated, assumed 2005	EMR vs. paper for net benefit/cost	Hardware, software, implementation, training, including discounts and incentives	Tangible cost savings in reduced staffing for managing paper chart, transcription and improved claims	Comparison of costs and benefits, return on investment	Costs of $213,083 and benefits of $657,500, return on investment of 308%	+ Excluded lost productivity from implementation and intangible benefits
Wang et al. (2003) USA	5 years not stated, assumed 2002-2006	EMR vs. paper for net benefit/cost	System and induced costs from historical records, experts and literature, discounted at 5% in 2002 US dollars	Estimated cost savings, increased revenues (payer independent, capitation, fee for service), discounted at 5% in 2002 US dollars	Net benefit per provider, with 1, 2 and 5-way sensitivity analysis	Net benefit $86,400 per provider over 5 years	+ based on proportion of patient capitation

Summary of Economic Evaluation Findings

Study	Time Frame	Options	Cost	Outcome	Comparison	Results	Interpretation
Medication Management							
Fretheim et al. (2006) Norway	1-year, assumed 2002	Outreach visit, audit-feedback and computer reminders vs. usual care, cost-benefit	Software and technical support; outreach staffing, training, and travel; cost of drug, physician time, and per patient lab test and	Number, percent and cost of patients being prescribed thiazides vs. other meds, cost in 2002 US dollars with 4% discount on drug cost in year 2	Cost minimization for thiazides vs. others; cost-effectiveness on incremental cost per additional patient started on thiazides	Net annual cost was $53,395, cost per additional patient on thiazides was $454; net annual savings for national program modelled to $761,998	-/+ Intervention cost 2x cost savings in year-1 but could lead to modest savings in 2 years
Wu et al. (2007) Canada	10 years not stated, assumed 2004-2013	CPOE-meds vs. paper for ADE prevention, cost benefit	Historical system and provider workload costs used to estimate annual costs, discounted at 5% in 2004 US dollars	ADE rates estimated from literature and incidence at hospital; number of preventable ADEs	Incremental cost effectiveness ratio, with one-way sensitivity analysis	ICER= $12,700 per ADE averted; 32.3 ADEs averted per year or 261 ADEs over 10 years	+/- based on ADE rate, system and physician cost, and ability to reduce ADEs

Summary of Economic Evaluation Findings

Study	Time Frame	Options	Cost	Outcome	Comparison	Results	Interpretation
CPOE/CDSS							
Kaushal et al. (2006) USA	10 years 1993-2002	Single CPOE system for costs and benefits	Internal documents and staff interviews used to estimate capital and operational costs of CPOE, with discount at 7% and annualized values, adjusted for inflation to 2002 US dollars	Benefit data from literature, key informants, internal documents (e.g., operating budget), in savings from reduced ADEs, drug/lab use and improved nursing time use, with discount at 7%, annualized values, adjusted to 2002 US dollars, and 80% prospective reimbursement rate	Net benefit, net cumulative present value, and operating budget benefit from CPOE	Net benefit $16.7M ($2.2M annualized), operating budget benefits $21.3M, net cumulative present value $9.5M ($1.3M annualized), took 5 years to realize net benefit and >7 years to operating budget benefit	+ substantial savings, including operating budget savings over 10 years
Ohsfeldt et al. (2005) USA	5 years not stated, assumed 2001-2005	Statewide CIS-CPOE vs. current CIS for financial impacts	Existing IT infrastructure from survey and estimated CIS/CPOE initial-ongoing costs from vendor, used to simulate CIS/CPOE costs with quadratic interpolation by bed size, and with 5 year depreciation, 5 year borrowing horizon at 5% interest rate	CIS/CPOE cost estimates combined with hospital revenue and cost data for financial impact by hospital type, with 5 year borrowing horizon, 5 year depreciation at 5% interest rate, and through 3rd party payers and reduced errors	Comparison of simulated operating margins for 1st and 2nd year by hospital type, with sensitivity analysis	Operating margins for post-CPOE year-1 and year-2 showed decrease for all hospital types, and deficit for rural and critical access hospitals	+/- CPOE costs may not be financially feasible for small hospitals

Summary of Economic Evaluation Findings

Study	Time Frame	Options	Cost	Outcome	Comparison	Results	Interpretation
CPOE/CDSS							
Poley et al. (2007) Netherlands	1 year with 6 months each for pre-post intervention 2001	Pre vs. post CDSS implementation for lab blood test ordering - costs and impacts	Estimated intervention cost (development, installation) for CDSS and staff, lab request cost for staff, material and facility, based on tests per request and volume of requests	Actual lab request costs with volume and tests per request including break-even points	Cost comparison for intervention (CDSS) and lab requests, with t-test, break-even point and sensitivity analysis	Actual cost savings of €847 Euros per practice per 6 months, and break-even point in 5 months	+ no change in lab request volume but did reduce tests per request

Summary of Economic Evaluation Findings

Study	Time Frame	Options	Cost	Outcome	Comparison	Results	Interpretation
Institutional EHR							
Byrne et al. (2010) USA	7 years 2001-2007, shorter for some components 2003/4-2007	Single system with EHR, PACS, MAR and lab data exchange, compared to not having the tool or component for net value	IT acquisition, operation and maintenance costs estimated in literature, local budgets, contract documents and IT staff, adjusted to 2007 US dollars	Estimated impact from literature, experts, level of adoption and service use in 5 categories: freed space and reduced workload, expense, utilization and redundancy in 2007 US dollars	Modelled net value based on annual and cumulative costs and potential benefits	Potential cumulative benefit $3.09B net investment cost by 7th year	+ spent more on IT than private health care sector but achieved higher IT adoption level and quality of care
Schmitt and Wofford (2002) USA	7 years 2000-2007	EMR with CPOE for radiology, pharmacy and lab; outpatient documentation	Hardware, software, implementation, security devices, imaging, technical support, at 10% discount rate	Staffing time in order processing, manage information/charts, and documentation; revenue from charge capture and claims; ADE rate	Estimated cost benefit and net present value	Projected net benefits of $49,519,094 or net present value of $31,360,953	+ Break-even point in year-2 with net benefits in year-3

Summary of Economic Evaluation Findings

Study	Time Frame	Options	Cost	Outcome	Comparison	Results	Interpretation
Disease Management							
O'Reilly et al. (2012) Canada	40 year horizon 2010-2049	Intervention vs. control groups compared for treatment effects	Historical immediate (1 year) and long-term (10 years) costs of 7 diabetes complications, and CDSS implementation costs, with 5% discount rates for costs and effects, in 2010 Canadian dollars	Relative risk reduction of complications from year-1 treatment; ICER from net cost of implementing CDSS, cost of treating complications and effectiveness over patient's lifetime	QALYs from year-1 treatment, relative risk reduction in complications, and incremental cost effectiveness ratio, with one-way sensitivity analysis	From year-1 treatment: incremental mean lifetime cost per patient $1,912, 14% relative risk reduction amputation QALY=0.0117, ICER=$156,970 per life year gained, and $160,845 per QALY gained	+ Modest improvement in short-term risk factors and moderate improvement in long-term health outcomes, but costly CDSS intervention

Summary of Economic Evaluation Findings

Study	Time Frame	Options	Cost	Outcome	Comparison	Results	Interpretation
Immunization Info System							
McKenna et al. (2002) USA	2 years 1998-1999	Immunization information system vs. paper for cost savings	Staff salary, budget and financial reports on development/operating costs for system and staff; observations and interviews for data entry time in year-1, with amortization of investments per child	Projected annual costs for year-2 with development and operating costs, with amortization of total investments and costs per child	Cost comparison for net savings	Actual cost savings in year-1 = $26,768, projected savings in year-2 = $689,403	+ if providers use registries and keep data current

Summary of Economic Evaluation Findings

Study	Time Frame	Options	Cost	Outcome	Comparison	Results	Interpretation
Documentation							
Kopach et al. (2005) Canada	4 years 2003-2006	Automated vs. traditional medical documentation for discharge note completion time	Current costs were system maintenance and staff for transcription; estimated automation costs were system, software and staff, cost and note volume discounted at 3%, system depreciated in 4 years, in 2003 Canadian dollars	Historical discharge notes used to calculate mean delay documentation time. Estimated reduction from literature, automation features and historical discharge notes, note volume discounted at 3%, discharge volume increase 1%	Incremental cost-effectiveness ratio, with one-way sensitivity analysis	ICER= $0.331 per day in average discharge note completion time	+/- automation more expensive but cost-effective based on physician utilization volume and length of study

Summary of Economic Evaluation Findings

Study	Time Frame	Options	Cost	Outcome	Comparison	Results	Interpretation
HIE							
Walker et al. (2005) USA	10 years not stated, assumed 2003-2012	Three HIE levels: transportable, organizable, interpretable, for net cost/benefit comparison	Interface and provider system costs from acquisition to ongoing maintenance based on literature and experts, in 2003 US dollars; only provider-payer costs were amortized over 3 years	Estimated cost savings from improved lab and imaging testing, prescribing, chart access, referral, provider-payer transactions, public health reporting, in 2003 US dollars	Net benefit, with sensitivity analysis	CBR= $77.8B at level 4, $23.9B at level 3, $21.6B at level 2 by year-11 steady state	+ based on lab and radiology tests

Summary of Economic Evaluation Findings

Study	Time Frame	Options	Cost	Outcome	Comparison	Results	Interpretation
HIE							

Institutional=Institutional Information System; EMR=Primary Care Electronic Medical Record; Documentation=Clinical Documentation System; Disease=Disease Management System; Immunization=Immunization Information System; Medication=Medication Management System; CDSS=Clinical Decision Support System; CPOE=Computerized Provider Order Entry; HIE=Health Information Exchange

ADE: adverse drug event
ADL: activities of daily living
AHA: American Hospital
CBR: cost-benefit ratio
CDR: clinical data repository
CDS: computerized decision support
CIS: clinical information system
DRG: diagnosis-related group
EHR: electronic health record
eMAR: electronic medication administration record
FTE: full-time equivalent
HIMSS: Health Information and Management Systems Society
ICER: incremental cost-effectiveness ratio
IT: information technology
LOS: length of stay
MAR: medication administration record
MDS: minimum data set
N/A: not applicable
PACS: picture archiving and communication system
PCIS: patient care information system
QALY: quality-adjusted life years
US: United States

References cited in Appendix

Block B. (2008). How we improved our practice and our bottom line with a new EMR system. *Family Practice Management, 15*(7), 25–30.

Byrne, C. M., Mercincavage, L. M., Pan, E. C., Vincent, A. G., Johnston, D. S., & Middleton, B. (2010). The value from investments in health information technology at the U.S. Department of Veterans Affairs. *Health Affairs (Millwood), 29*(4), 629–638.

Fretheim, A., Aaserud, M., & Oxman, A. D. (2006). Rational prescribing in primary care (RaPP): economic evaluation of an intervention to improve professional practice. *PLoS Medicine, 3*(6), e216.

Grieger, D. L., Cohen, S. H., & Krusch, D. A. (2007). A pilot study to document the return on investment for implementing an ambulatory electronic health record at an academic medical center. *Journal of the American College of Surgeons, 205*(1), 89–96.

Karnon, J., McIntosh, A., Dean, J., Bath, P., Hutchinson, A., Oakley, J., … Tappenden, P. (2008). Modelling the expected net benefits of interventions to reduce the burden of medication errors. *Journal of Health Services Research and Policy, 13*(2), 85–91.

Kaushal, R., Jha, A. K., Franz, C., Glaser, J., Shetty, K. D., Jaggi, T., … Bates, D. W. (2006). Return on investment for a computerized physician order entry system. *Journal of American Medical Informatics Association, 13*(3), 261–266.

Kopach, R., Geiger, G., & Ungar, W. J. (2005). Cost-effectiveness analysis of medical documentation alternatives. *International Journal of Technology Assessment in Health Care, 21*(1), 126–131.

Kumar, S., & Bauer, K. (2011). The business case for implementing electronic health records in primary care settings in the United States. *Journal of Revenue and Pricing Management, 10*(2), 119–131.

McKenna, V. B., Sager, A., Gunn, J. E., Tormey, P., & Barry, M. A. (2002). Immunization registries: costs and savings. *Public Health Reports, 117*(4), 386–392.

Miller, R. H., West, C., Brown, T. M., Sim, I., & Ganchoff, C. (2005). The value of electronic health records in solo or small group practices. *Health Affairs (Millwood), 24*(5), 1127–1137.

Ohsfeldt, R. L., Ward, M. M., Schneider, J. E., Jaana, M., Miller, T. R., Lei, Y., & Wakefield, D. S. (2005). Implementation of hospital computerized physician order entry systems in a rural state: Feasibility and financial impact. *Journal of American Medical Informatics Association, 12*(1), 20–27.

O'Reilly, D., Holbrook, A., Blackhouse, G., Troyan, S., & Goeree, R. (2012). Cost-effectiveness of a shared computerized decision support system for diabetes linked to electronic medical records. *Journal of American Medical Informatics Association, 19*(3), 341–345.

Patil, M., Puri, L., & Gonzalez, C. M. (2008). Productivity and cost implications of implementing electronic medical records into an ambulatory surgical subspecialty clinic. *Urology, 71*(2), 173–177.

Poley, M. J., Edelenbos, K. I., Mosseveld, M., van Wijk, M. A., de Bakker, D. H., van der Lei, J., & Rutten-van Mölken, M. P. (2007). Cost consequences of implementing an electronic decision support system for ordering laboratory tests in primary care: evidence from a controlled prospective study in the Netherlands. *Clinical Chemistry, 53*(2), 213–219.

Schmitt, K. F., & Wofford, D. A. (2002). Financial analysis projects clear returns from electronic medical records. *Healthcare Financial Management, 56*(1), 52–57.

Simon, S. J., & Simon, S. J. (2006). An examination of the financial feasibility of electronic medical records (EMRS): a case study of tangible and intangible benefits. *International Journal of Electronic Healthcare, 2*(2), 185–200.

Walker, J., Pan, E., Johnston, D., Adler-Milstein, J., Bates, D. W., & Middleton, B. (2005). The value of health care information exchange and interoperability. *Health Affairs (Millwood), Suppl Web Exclusives,* W5-10–W5-18.

Wang, S. J., Middleton, B., Prosser, L. A., Bardon, C. G., Spurr, C. D., Carchidi, P. J., … Bates, D. W. (2003). A cost-benefit analysis of electronic medical records in primary care. *The American Journal of Medicine, 114*, 397–403.

Wu, R. C., Laporte, A., & Ungar, W. J. (2007). Cost effectiveness of an electronic medication ordering and administration system in reducing adverse drug events. *Journal of Evaluation in Clinical Practice, 13*(3), 440–448.

Chapter 19

Evaluation of eHealth System Usability and Safety

Morgan Price, Jens Weber, Paule Bellwood, Simon Diemert, Ryan Habibi

19.1 Introduction

Usability and safety are two types of non-functional requirements[1] or quality attributes of a system. Both are increasingly important in health information and communication technology (ICT) systems as they become more integrated into care processes from primary care to the intensive care unit (ICU). Usability and safety are emergent properties of systems, not a property of any particular device such as a piece of computer software. Thus, both should be considered in the context of the *sociotechnical system* of which they are parts. In this chapter, we consider both usability and safety, as we feel they can and should be related.

19.2 Definitions

Sociotechnical systems comprise technology (software and hardware), actors (such as patients, providers, caregivers, friends, and administrators), physical spaces, and the policies that interact, in our case, to support health and wellness. A sociotechnical system in primary care may be a complex web of actors which make up a patient's circle of care and related technologies. For example: a physician office with physicians, nurses, staff and an electronic record; a pharmacy

1 Non-functional requirements are requirements that do not describe a specific behaviour of a system but rather a requirement that describes how a system is judged to be and is architected into the system as a whole. There are several types of non-functional requirements including: usability, safety, availability, scalability, effectiveness, and testability.

with pharmacists and pharmacy technicians all working through an information system; a person working with their physical trainer who starts using a pedometer and some mobile Health apps to track weight, activity and diet.

Usability is the ease with which a system can be used by the intended actors to achieve specified goals. It also includes a system's learnability. Usability considers satisfaction, efficiency, effectiveness, and context of use (see ISO standard 9241-11). Usability is deeper than the look and feel of a system or user satisfaction; it also includes how a system works in context to complete work or manage workflows, and how well that fits with the needs of users. Usability includes how easy the system is to learn for users and how quickly users can relearn the tool if it is upgraded or if it is not used for a period of time. Finally, usability can positively or negatively impact safety.

Safety is "freedom from those conditions that can cause death, injury, occupational illness, damage to or loss of equipment or property, or damage to the environment" (United States Department of Defense, 2012). Devices (or components of devices) are referred to as *safety-critical* if they are essential for the safe operations of systems of which they are a part (i.e., their failure alone could result in death, injury, or loss). Otherwise, devices are referred to as *safety-sensitive* if they contribute to safety-critical functions.

Depending on their respective impacts on safety, devices used in eHealth systems may be subject to different levels of mandatory regulation, evaluation, and certification, which may include pre-market evaluation as well as post-market surveillance (Weber-Jahnke & Mason-Blakley, 2012). In practice, however, the classification with respect to their safety impact of many of the devices used in eHealth systems has been challenging. Regulators have struggled to develop a balanced framework for eHealth system evaluation and control. There are two main reasons for these problems: firstly, eHealth devices such as Electronic Medical Records (EMRs) are often complex aggregates of many diverse functions with different criticality; and secondly, systems these devices are integrated into are highly diverse and variable, and by necessity may not be as expected by the device manufacturer.

There are frequent and subtle interactions between the usability and the safety of eHealth systems (see Figure 19.1), which evaluators need to be aware of. In some cases, there may be trade-offs between these two types of requirements. Safety mechanisms may decrease the perceived usability of a system (e.g., where users are required to click on medication alerts while prescribing). Usability enhancements may decrease the safety of a system (e.g., where users are given the opportunity to skip or automate certain tasks). In other cases, increased usability may actually lead to increased safety (e.g., a clean, uncluttered user interface may reduce cognitive load and help prevent medical errors).

The above considerations emphasize the importance of considering larger systems while designing, modelling, and evaluating eHealth devices where sociotechnical aspects of both usability and safety interact (Borychi & Kushniruk,

2010). Thus, it is important to consider safety and usability and their interactions while evaluating any given system.

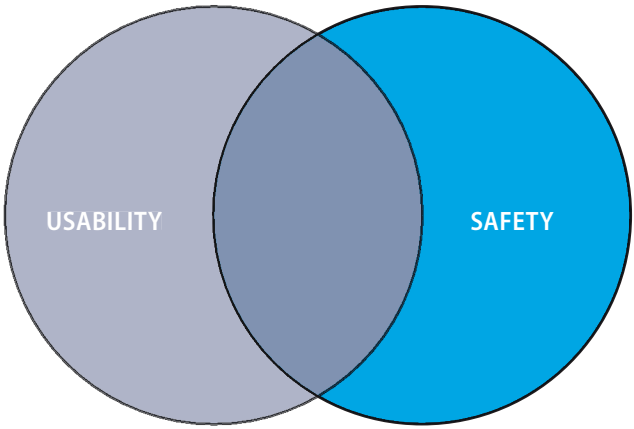


Figure 19.1. Usability and safety requirements often overlap and there is value in considering both.

19.3 When to Evaluate

The importance of evaluating the usability of eHealth systems has been highlighted for almost two decades (Friedman & Wyatt, 1997). Initial usability evaluation in eHealth focused on post-implementation evaluations; however, it has become increasingly evident that these systems should be evaluated sooner in their life cycles, starting from the project planning stages through design and implementation (Kushniruk, 2002; Kushniruk & Patel, 2004; Marcilly, Kushniruk, Beuscart-Zephir, & Borycki, 2015). Conversely, initial safety evaluation efforts of eHealth systems have focused on pre-implementation evaluations, while more recent evidence indicates the insufficiency of this approach and the need for additional post-implementation evaluations.

Ideally, evaluation of usability and safety of eHealth systems should occur throughout their life cycle — during conception, design, development, deployment, adoption, and ongoing evolution. While evaluation should be considered throughout the life cycle, the methods and focus of the evaluation may change over time. Current evaluations of eHealth systems are aimed at evaluating the technology in early stages of design to make informed design decisions and reduce risks; additionally, evaluating during implementation and post-deployment to assess the impact of a system and improve future system revisions (Marcilly et al., 2015). Earlier evaluation during design and/or procurement of systems is considerably less expensive than trying to change existing tools and processes post-implementation.

Choosing not only the proper methods to evaluate eHealth systems throughout their life cycles but being aware of the contexts in which to evaluate these

systems is essential (Kuziemsky & Kushniruk, 2014, 2015). For example, when designing a system, one can employ usability testing and safety inspection methods on low fidelity prototypes and workflow designs, respectively. As a system is deployed, observational studies are very useful to understand how it is used in practice and one may see surprising workflows, workarounds, and unintended consequences. Thus, these different methods help support decision-making with regard to the eHealth system, how it is designed, configured, and implemented.

19.4 Usability Methods

There are many methods for assessing and improving the usability of systems. It is helpful to broadly categorize these methods first, before providing a few examples. Usability methods can be broadly categorized into *inspection methods* and *testing methods*. Usability inspection methods, as a group, are expert-driven assessments of a design or product's usability. They do not involve users. Usability testing methods, by contrast, engage real-world users — potential or expected users — to explore user interfaces, often completing important or common tasks within the system that test both the user interface and user experience.

Both types of usability methods can vary in their focus. For example, they can be very granular, focusing on an individual's interaction with the eHealth application, or they can focus on the broader interactions between actors in a group. Table 19.1 provides some examples in each category. A system's usability can be evaluated in different settings, including real (i.e., in-situ) or simulated environments (i.e., clinical simulations in a usability lab). Using clinical simulations for usability evaluations often results in higher evaluation fidelity (Borycki, Kushniruk, Anderson, & Anderson, 2010; Li et al., 2012).

Table 19.1
Usability Methods Categorized by Type and Focus

	Individual Focus	Group Focus
Inspection	• Cognitive Task Analysis • Heuristic Inspection	• Distributed Task Analysis
Testing	• Think Aloud User Testing	• Observational Studies • Contextual Analysis

- *Cognitive Task Analysis* is a form of expert inspection that focuses on the cognitive needs of an individual user (in a particular role) as they complete tasks. Cognitive Task Analysis is well suited for eHealth systems; much of healthcare is focused on the cognitively intensive tasks of collection and synthesizing patient information for diagnoses and managing treatment.

- *Think Aloud* is a common form of usability testing where individual users are asked to use an application and encouraged to speak their mind while completing tasks. By thinking aloud in the moment, the designers are able to capture usability challenges that might not otherwise be remembered by the user in follow-up interviews. Multiple users are asked to individually complete a set of tasks in the application, typically while being recorded. The analyst then reviews the session (or their notes) to highlight usability challenges in using the system to complete the tasks. The findings across the multiple test sessions are then synthesized into design recommendations that can be implemented and retested.

- *Distributed Task Analysis* builds on the theory of Distributed Cognition (Hutchins, 1995) and is a model that expands the concept of cognition outside of the mind to groups of actors (both human and technical). Understanding how a patient is kept alive in a trauma in an emergency or during surgery are two examples where a distributed task analysis would be helpful as there are many actors working together in parallel. Like cognitive task analysis, distributed task analysis is an inspection method; however, the scope is typically larger, considering how a process unfolds and how groups of actors (and in this case eHealth tools) work together to come to decisions and complete actions.

- *Observational Studies* place the analyst within an environment to observe the context of work. There are several approaches to observational studies, with varying focus, methods for recording observations (from note taking to digital recording of audio and video), and duration. Observational studies permit better understanding of the interactions between the technology and the interdependent workflow between actors (people, patients, physicians, nurses, etc.). Observations can take place at single or multiple locations and may focus on care flows of single patients through the healthcare system, or can be team focused, observing how a ward or department might work.

19.5 Safety Methods

As highlighted previously, the quality attribute of safety is often linked to that of usability. Consequently, the usability evaluation methods as characterized above may also be helpful for identifying safety-related concerns, in particular when it comes to safety concerns related to human factors and human-computer interaction. A variety of methods have been developed for evaluating sys-

tems for safety concerns. What follows is a description of four prominent methods for evaluating system safety.

1 *System Theoretic Accident Model and Processes* (STAMP) is a method that been developed in the systems engineering context and seeks to model systems as interacting control loops (Leveson, 2012). This method defines a taxonomy of different classes of safety-sensitive errors to be considered in the analysis. Safety is assured by putting in place (and enforcing) constraints on the behaviour of components in the system-theoretic model. STAMP can be used at different stages of the life cycle from requirements to (and after) deployment. STAMP provides systematic methods for retrospective accident analysis, that is, for identifying missing safety constraints that may have contributed to accidents or near misses, as well as methods for prospective design of safe systems. Figure 19.2 illustrates the concept of using control loops as a system-theoretic model for representing EMR-based care processes.

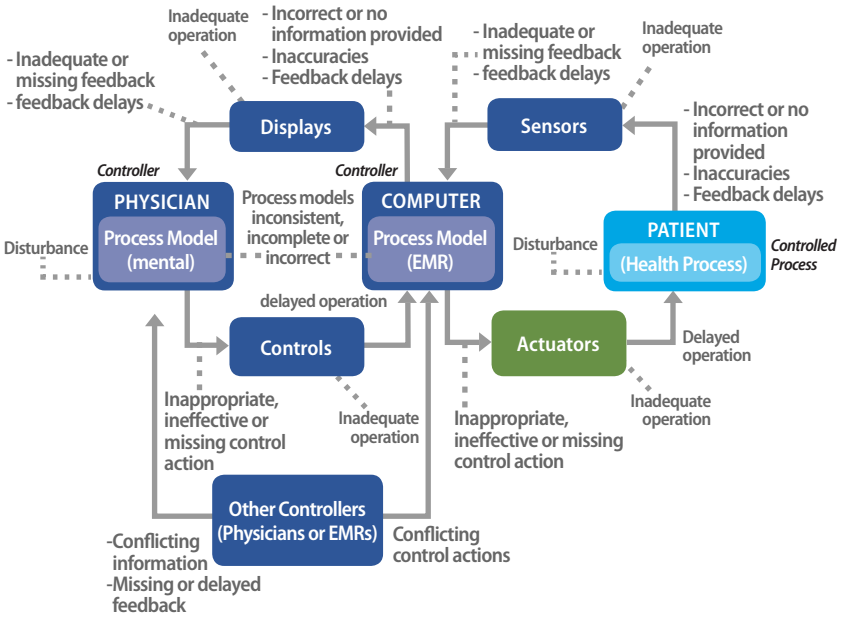


Figure 19.2. STAMP applied to EMR systems.

Note. From "On the safety of electronic medical records," by J. Weber-Jahnke and F. Mason-Blakley, 2012, *First International Symposium, Foundations of Health Informatics Engineering and Systems (FHIES)*, p. 186. Copyright 2012 by Springer. Reprinted with permission.

2 *Failure Modes and Effects Analysis* (FMEA) is a method developed by the safety engineering community, which has also been adapted to healthcare as Healthcare FMEA (HFMEA), and has been used by the U.S. Department of Veterans Affairs (DeRosier, Stalhandske, Bagian, & Nudell, 2002). The method is based on a process model describing the relevant workflows within a particular system. It systematically identifies potential failure modes associated with the system's components and determines possible effects of these failures. Failures are assigned criticality scores and are ranked accordingly. Control measures are developed to mitigate accidents that could result from the most critical failure modes. HFMEA can be used early in the design of new systems or processes and also much later as the sociotechnical systems evolve with time and use.

3 *Fault Tree Analysis* (FTA) is a deductive method that starts by assuming safety faults and successively seeks to identify conditions under which system components could lead to these faults (Xing & Amari, 2008). An example of a system fault in the healthcare domain could be *patient has an adverse reaction to a medication*. Conditions which could lead to such a fault could include malfunctions of the clinical decision support system (for showing drug allergy alerts), malfunction of the communication system between the EMR and pharmacy, missing or incongruent data in the EMR about the patient (allergies, other active medications, etc.), or other factors. FTA successively analyzes potential causes for safety faults in a hierarchical (tree-like) structure; this is a deductive approach and complementary to FMEA, which is inductive in nature. By contrast, FMEA starts from system components, their potential failure modes and focuses on determining possible faults that could result from them.

4 *Hazard and Operability* (HAZOP) is another process-based safety evaluation method, which was originally developed in the design of industrial chemical plants, but has since been used for computer-based systems (Dunjó, Fthenakis, Vílchez, & Arnaldos, 2010). HAZOP relies on a disciplined, systematic process of using guidewords to discover potential unintentional hazardous consequences of process deviations. Typical HAZOP guidewords include "no", "more", "less", "as well as", "reverse", etc. These guidewords are applied to actions modelled in the process under investigation to identify possible process deviations and their (potentially safety-relevant) consequences.

19.6 Selected Case Study Examples

The following two examples have been selected because they both have aspects of usability and safety. The first example is primarily safety focused, examining a commonly cited case study of a computer-based physician order entry (CPOE) system. The second example illustrates how usability design standards were developed in order to improve overall safety of eHealth in the United Kingdom's National Health Service (NHS).

19.6.1 Safety Case Study: A Technology-induced Medication Error

The first case study involves a CPOE system deployed at the New York Presbyterian Hospital. Horsky, Kuperman, and Patel (2005) analyzed the factors that led to a technology-induced medical accident, while Weber-Jahnke and Mason-Blakley (2012) provided a further systematic analysis using a STAMP. In this incident, an elderly patient was admitted to the hospital and received a significant overdose of Potassium Chloride (KCl) over a period of two days, involving multiple medication orders by multiple providers. Notably, no single event can be pinpointed as the root cause for the accident and the CPOE device functioned as intended by the manufacturer. Rather, the accident was the result of a number of factors that in combination resulted in the harmful outcome.

The following is a series of significant events leading to the harmful outcome (i.e., an accident):

1. On Saturday, Provider A reviews the results of a lab test and finds the patient hypokalemic (deficient in bloodstream potassium).

2. Provider A orders a KCl bolus injection using the CPOE.

    a. Provider A notices that the patient has an existing drip line and decides to use the line instead of an injection.

    b. Provider A enters a new drip line order and intends to cancel the injection order.

    c. However, Provider A inadvertently cancels a different (outdated) injection order, which had been entered by a different provider two days prior.

3. Provider A is notified by the pharmacy because the dose for the drip order exceeds the hospital's maximum dose policy.

4. Provider A enters a new drip order but fails to enter it correctly (a maximum volume of 1L was entered but in the wrong input field, namely the "comment" field).

a. Provider A enters this information in the "comment" field as free text but fails to enter it in the structured part of the CPOE input form.

5 The KCl fluid continues to be administered for 36 hours, in addition to the initial bolus injection that ran to completion.

6 On Sunday morning, Provider B takes over the case and checks the patient's KCl level based on the most recent lab test (which was still from Saturday).

7 Not realizing that the patient's initial hypokalemic state had already been acted upon, Provider B orders two additional KCl injections.

8 On Monday morning a KCl laboratory test found the patient to be severely hyperkalemic. The patient was treated immediately for hyperkalemia.

This case study highlights several aspects related to usability, safety, and the interaction between these two system quality attributes:

A. The failure to specify an effective stop date / maximum volume for Provider A's drip order is a direct result of a usability problem. The CPOE input form allowed the provider to make free text comments on the order, but these comments were not seen as instructions by the medication-administering nurses.

B. The failure of Provider B to realize the patient's hypokalemic state is a clear system (safety) design problem. The device could have been designed to relate ordered interventions to out-of-range test results, and make providers aware of the fact that test results had already been acted on.

C. The failure of Provider A to cancel the right order cannot clearly be categorized as a sole usability or safety problem, respectively. Rather, it relates to both aspects. On one hand, the device could have made it easier to distinguish old (and new) orders from orders submitted by other providers (and in the past). On the other hand, a more effective design of the CPOE device could have detected an overdose violation based on the consideration of multiple orders rather than based only on the consideration of each order separately.

Usability and safety evaluation studies may have prevented or mitigated the above accident. For example, Think Aloud user testing with providers may have indicated that providers tend to use the "comment" field of the CPOE device to specify volume limits, while administering nurses would disregard that field (see point A above). Safety evaluation methods may have prevented point B. For example, the application of HAZOP guidewords like "as well as" on the order entry process step (after the lab review step) may have revealed the hazard of prescribing interventions more than once as a reaction to a specific lab test. Ideally, proper design mitigation would have flagged the out-of-range lab test as "already acted upon" in the EMR. Finally, usability or safety evaluation methods could have mitigated point C above. For example, cancelling the wrong medication order is a clear failure mode of the ordering system (FMEA), which could be mitigated by checking whether the cancelled order is current, or has already been administered in the past. Moreover, HAZOP guidewords could have identified the hazard of medication overdoses due to two or more concurrent medication orders of the same substance.

19.6.2 Usability Case Study: Common User Interface

The *Common User Interface* project (CUI) was an attempt to create a safer and more usable eHealth user interface by defining a standard across multiple clinical information systems that would be consistent for users. This project was undertaken as a joint effort between the U.K.'s National Health Service (NHS) and Microsoft. Safety through improved user interface design was a key consideration. As part of a larger project, CUI set about to create *design guidances* that presented a standard (common) user interface approach for aspects of eHealth tools that would better support care. Further, this would support clinicians who were moving between different eHealth systems. The CUI design guidances were published and cover a range of topics within the following:

- Patient identification
- Medications management
- Clinical notes
- Terminology
- Navigation
- Abbreviation
- Decision support

Each design guidance is an extensive document that addresses a component of one of the topics above. For example, as part of the medications management guidelines, there are detailed documents for "drug administration", "medication line", and "medication list" among others that help developers with specific information on how to (and how not to) implement the user interface. The design guidance documents were developed in a manner compliant with the Clinical Safety Management System defined by the NHS. Furthermore, the guidelines include the rationale for the recommendations (and associated evidence).

For example, the medication line design guideline (v2.0.0)[2] carefully describes how a medication should be displayed. It includes specific recommendations for display of generic names, brand names, strength, dose, route, and frequency. These include rationale for font styles, spacing, and units that make information easier to read, to comprehend, and reduce the risk for misinterpretation. Figure 19.3 demonstrates CUI guidances such as: "generic medication name must be displayed in bold"; "dose must be clearly labelled"; "acronyms should not be used when displaying the medication instructions"; and "instructions should not be truncated but all instructions must be shown, with wrapping if necessary" (note oxycodone uses three lines).

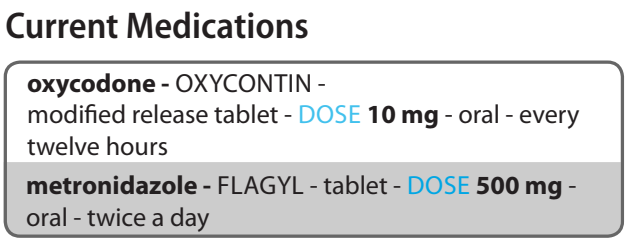


Figure 19.3. An example of medication display following CUI design guidance.

The Microsoft Health Patient Journey Demonstrator was built to demonstrate how CUI guidances could be implemented on a Microsoft platform to display health information in a health information system (Disse, 2008). This example, showing how CUI could be applied to primary care, secondary care, as well as administrative clinical interfaces, has attracted attention from various communities due to its applicability and as a solution to provide a standardized approach to clinical user interfaces. The CUI design guidances are freely available[3]. Microsoft© also provides some free example software controls under the Microsoft Public License.

CUI was an impressive effort and reviewing many of the guidelines in these design guidances provides a wealth of information on how to and how not to

2 http://systems.hscic.gov.uk/data/cui/uig/medline.pdf
3 http://systems.hscic.gov.uk/data/cui/uig

design user interfaces in the health domain. However, CUI only covered a small number of areas and the project has not continued. The knowledge that was generated is freely available at mscui.org and through the NHS.

19.7 Summary

Usability and Safety are increasingly being acknowledged as necessary components for the success of eHealth. However, achieving safe and usable systems remains challenging. This may be because it is often unclear how to measure these quality attributes. Further, as systems are deployed and adopted, it becomes harder and more costly to make large changes. This is especially the case as eHealth tools are being increasingly integrated into care processes across the circle of care, and as people and providers use an increasing range of tools, apps and health records to manage care.

A single, large "safety review" or "usability inspection" is less likely to have a long-lasting impact. Instead organizations should focus on embedding usability and safety in their culture and process. Thus, we encourage that safety and usability engineering should occur throughout the life cycle of eHealth tools from requirements and procurement to ongoing evaluation and improvement. In this chapter we have highlighted a few methods for evaluating safety and usability. It is likely more feasible to build on existing work, such as the CUI project, and use multiple methods to triangulate findings across small evaluation projects than it is to attempt a large, comprehensive study with a single method; multiple methods complement each other.

Policy-makers, funding programs, and health organizations should explicitly embed safety and usability engineering into the operational eHealth processes. There is increasing need for both usability and safety engineers in health as eHealth systems are being, and continue to become, broadly adopted.

References

Borycki, E., & Kushniruk, A. (2010). Towards an integrative cognitive-socio-technical approach in health informatics: Analyzing technology-induced error involving health information systems to improve patient safety. *The Open Medical Informatics Journal, 4*, 181–187. doi: 10.2174/1874431101004010181

Borycki, E., Kushniruk, A. W., Anderson, J., & Anderson, M. (2010). *Designing and integrating clinical and computer-based simulations in health informatics: From real-world to virtual reality*. Vukovar, Croatia: In-Tech.

DeRosier, J., Stalhandske, E., Bagian, J. P., & Nudell, T. (2002). Using health care failure mode and effect analysis: The VA National Center for Patient

Safety's prospective risk analysis system. *The Joint Commission Journal on Quality Improvement, 28*(5), 248–267.

Disse, K. (2008). Microsoft health patient journey demonstrator. *Informatics in Primary Care, 16*(4), 297–302.

Dunjó, J., Fthenakis, V., Vílchez, J. A., & Arnaldos, J. (2010). Hazard and operability (HAZOP) analysis. A literature review. *Journal of Hazardous Materials, 173*(1), 19–32. doi: 10.1016/j.jhazmat.2009.08.076

Friedman, C. P., & Wyatt, J. (1997). *Evaluation methods in medical informatics*. New York: Springer.

Horsky, J., Kuperman, G. J., & Patel, V. L. (2005). Comprehensive analysis of a medication dosing error related to CPOE. *Journal of the American Medical Informatics Association, 12*(4), 377–382. doi: 10.1197/jamia.M1740

Hutchins, E. (1995). How a cockpit remembers its speeds. *Cognitive Science, 19*(3), 265–288.

Kushniruk, A. (2002). Evaluation in the design of health information systems: Application of approaches emerging from usability engineering. *Computers in Biology and Medicine, 32*(3), 141–149. doi: 10.1016/S0010-4825(02)00011-2

Kushniruk, A. W., & Patel, V. L. (2004). Cognitive and usability engineering methods for the evaluation of clinical information systems. *Journal of Biomedical Informatics, 37*(1), 56–76. doi: 10.1016/j.jbi.2004.01.003

Kuziemsky, C. E., & Kushniruk, A. (2014). Context mediated usability testing. *Studies in Health Technology and Informatics, 205*, 905–909.

Kuziemsky, C., & Kushniruk, A. (2015). A framework for contexual design and evaluation of health information technology. *Studies in Health Technology and Informatics, 210*, 20–24.

Leveson, N. (2012). *Engineering a safer world: Systems thinking applied to safety*. Cambridge, MA: MIT Press.

Li, A. C., Kannry, J. L., Kushniruk, A., Chrimes, D., McGinn, T. G., Edonyabo, D., & Mann, D. M. (2012). Integrating usability testing and think-aloud protocol analysis with "near-live" clinical simulations in evaluating clinical decision support. *International Journal of Medical Informatics, 81*(11), 761–772. doi: 10.1016/j.ijmedinf.2012.02.009

Marcilly, R., Kushniruk, A. W., Beuscart-Zephir, M., & Borycki, E. M. (2015). Insights and limits of usability evaluation methods along the health information technology lifecycle. *Studies in Health Technology and Informatics, 210*, 115–119.

United States Department of Defense. (2012). Standard practice for system safety: MIL-STD-882E. Retrieved from http://www.system-safety.org/Documents/MIL-STD-882E.pdf

Weber-Jahnke, J., & Mason-Blakley, F. (2012). On the safety of electronic medical records. In Z. Liu & A. Wassyng (Eds.), *Foundations of Health Informatics Engineering and Systems: First international symposium, FHIES 2011* (pp. 177–194). Berlin: Springer.

Xing, L., & Amari, S. V. (2008). Fault tree analysis. In K. B Misra (Ed.), *Handbook of performability engineering* (pp. 595–620). London: Springer.

Chapter 20

Evaluation of eHealth Adoption in Healthcare Organizations

Jim Warren, Yulong Gu

20.1 Introduction

Healthcare innovations, including eHealth technologies, aim to support faster, more reliable and more transparent healthcare services. These technologies may facilitate the design and delivery of high-quality healthcare, improved patient outcomes and patient safety, and further generation of innovation in healthcare processes (Chaudhry et al., 2006; Finkelstein et al., 2012; Lau, Kuziemsky, Price, & Gardner, 2010). However, adoption of eHealth technologies in healthcare organizations involves complex sociotechnical issues and often fails (Kaplan & Harris-Salamone, 2009). Adopters of eHealth technologies face challenges such as the complexity inherent to the healthcare services context and a multitude of risk factors in the eHealth development/procurement and implementation processes. These challenges need to be understood and promptly addressed to support successful implementation and sustained use of health innovations. As such, one fundamental starting goal for any process of eHealth evaluation is to evaluate adoption, particularly in terms of the uptake of the technology by the intended end users.

In chapter 6 we categorized uptake (or simply "use") under the "product" dimension of our Health Information Technology (IT) Evaluation Framework as a component of usability. While this is valid, uptake is an essential step on the pathway to any and all progress on the "impact" dimension of the framework (i.e., for improvement in work and communication patterns, organizational culture, safety and quality of healthcare, or overall effectiveness). This is well illustrated by a study of clinical decision support effectiveness for chronic condition management that was undertaken in the context of general practice in the

United Kingdom (U.K.). The system showed no impact whatsoever on process of care or patient outcomes, while noting that usage was low (Eccles et al., 2002). Further in-depth investigation with users identified barriers to use that included concerns about the timing of the guideline trigger, ease of use and helpfulness of content, as well as problems in the delivery of training (Rousseau, McColl, Newton, Grimshaw, & Eccles, 2003). The lack of overall impact becomes unsurprising once the story with respect to barriers to adoption is understood.

Health IT adoption has received considerable attention in recent years. For instance, the Healthcare Information and Management Systems Society (HIMSS) Analytics Electronic Medical Record Adoption Model (EMRAM) ranks healthcare organization progress into one of seven stages based on the types of systems that are in place (HIMSS, 2015). In the government context, health IT adoption in the United States is being driven by financial incentives that are tied to achievement of a spectrum of specific "meaningful use" criteria (Marcotte et al., 2012). Similarly, the U.K. Quality and Outcomes Framework (QOF) provides substantial financial incentives to general practitioners (GPs; i.e., community-based family physicians) based on monitoring and management levels as automatically assessed through their practice electronic medical record systems (Lester & Campbell, 2010). It could be said that the aforementioned models are tied to ticking the boxes to achieve financial incentives (or "bragging rights"). At a more conceptual level, sophistication in health IT use can be broken down into technological support, information content, functional support, and IT management practices (Raymond & Paré, 1992), as well as extent of systems integration (Paré & Sicotte, 2001). A further dimension of IT sophistication, in terms of application domain, concerns administrative activities, patient care and clinical support; and in any of the above domains one can assess the range of computerized activities and system availability, as well as extent of use (Kitsiou, Manthou, Vlachopoulou, & Markos, 2010; Paré & Sicotte, 2001).

In this chapter, we illustrate the approach to evaluation of system adoption with two case studies that are based on electronic referral (eReferral) technologies. The United States National Library of Medicine has defined "referral and consultation" as "the practice of sending a patient to another program or practitioner for services or advice which the referring source is not prepared to provide" (National Library of Medicine, 2014), which implies a transfer of care. In the New Zealand (N.Z.) healthcare context, referral is most often from a GP to a specialist medical service. Moreover, the general practices in N.Z. tend to be private for-profit or charitable trust organizations (although supported by government subsidies). Individual general practice sites are small (for instance, they may be part of a strip mall) and are characterized as being situated in "the community" alongside other services including a community pharmacy and homecare nursing. Conversely, specialist services are provided in large part at public hospitals operated directly by District Health Boards; eReferral aims to use IT to bridge the communications gap between these two types of providers and their contrasting sites. While eReferral may simply replace a postal or fax pro-

cess with e-mail, more advanced IT offers opportunities for rich and rapid feedback that transforms the process. At the more extreme end, eReferral can merge into a portal-based "shared care" model that challenges the original concept of referral (Gu, Warren, & Orr, 2014).

Although the two case studies in this chapter are both situated in the New Zealand context of bridging community-hospital divides, we believe they can be generalized to any situation where IT is mediating healthcare communication across provider roles and sites. Moreover, these cases serve to illustrate contexts where providers can potentially work around uptake of the technology (e.g., sidestepping with phone and fax) and thus adoption is a valuable measure of success.

20.2 Selected Case Study Examples

In eHealth evaluation literature, both qualitative and quantitative methods have been used to measure a range of indicators on usability and outcome. The value of these evaluations is not limited to collecting robust evidence on the impact of eHealth innovations, which of course is important for measuring project success or supporting the decision-making process with regard to technology purchase and further rollout (or abandonment) of the technology adoption. The evaluation research can also provide substantial support to the technology development and implementation process. That is, if you evaluate early and often, learnings from evaluation can be used to improve the acceptability and effectiveness of the technology in its current implementation sites, as well as being fed into subsequent phases of implementation.

The following two examples of eHealth evaluation are introduced to demonstrate impact analysis-focused evaluation and Action Research-oriented evaluation, respectively.

20.2.1 Case Study One — Impact of an Electronic Referral System

This case describes a retrospective evaluation study of the impact of introducing an eReferral system that manages referrals from a community into public secondary healthcare services (Warren, White, Day, Gu, & Pollock, 2011). The eReferral system evaluated was introduced in 2007 to 30 referring general medical practices and 28 hospital-based secondary services at an N.Z. regional healthcare jurisdiction, Hutt Valley District Health Board (HVDHB). HVDHB serves a population of 150,000 and has one principal facility for provision of secondary services, the 260-bed Hutt Hospital.

By October 2007, eReferral to 28 services at Hutt Hospital — to all services but the Emergency Department — had been deployed across 25 general practices. A GP, or in some cases a practice nurse, creates a referral from within their electronic medical record system (in New Zealand, it is often called practice management system, or PMS) using PMS-based templates. The form is pre-populated with PMS data including the patient's demographics and medical history, which can be edited by the GP prior to submission. The GP referral is messaged as

Extensible Markup Language (XML) via the regional service where it is mapped to Health Level 7 (HL7) message format and sent on to the Integration Engine, which underlies the HVDHB's Clinical Workstation. The Integration Engine generates an acknowledgement back to the network confirming receipt of the referral, which is relayed back to the GP PMS.

Hospital staff can view the eReferral and process the eReferrals in a Clinical Workstation displaying in automated role-based work lists. It allows clinicians to communicate within the service administration to manage clinics, required tests, or follow-up with patients prior to the appointment. The referral management activity creates an automated sequence of process events through clinician triage (assignment of priority) and, if not declined at triage, regarding creation of a booking for a first specialist appointment (FSA). The relevant referral workflow is shown in Figure 20.1.

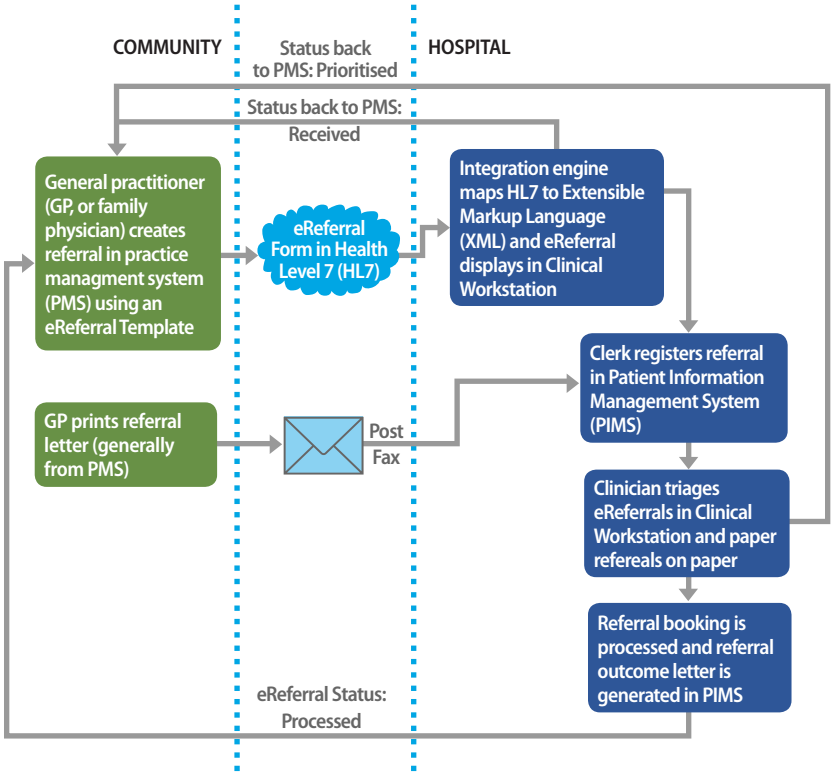


Figure 20.1. Referral workflow.

As of November 2010, there was no central referral management at Hutt Hospital; however, the general Outpatient Department with eight administration staff manages 15 services; the remainder receive and manage their own re-

ferrals. In the context of an N.Z. public hospital, services undertake clinical triage of referrals that assigns priority levels to them, including declining to service some requests (noting that private services are also available). Referral management within the hospital involves two concurrent systems: a Clinical Workstation and a Patient Information Management System (PIMS, providing general inpatient tracking, by a different vendor from the Clinical Workstation). All referrals (electronic or paper) are logged to the PIMS.

Based on a literature review, Hutt eReferral project business case and documentation review, and stakeholder feedback, the eReferral evaluation hypothesis was developed as: eReferral, if uptake is substantial and sustained, should result in more efficient (and thus timely), as well as more transparent, processing of referrals. To test this hypothesis, 33,958 transactional records from October 2007 to the end of October 2010 were collected from the eReferral database, as stored with the Clinical Workstation: 108,652 records of all GP referrals (electronic and paper) from the hospital PIMS were extracted from January 2004 to end October 2010. These data allowed examination of eReferral's impact, in terms of uptake (eReferral volume over time and proportion of referrals that are electronic) and changes in latency from letter date to triage at secondary services. The extracts, de-identified and using encrypted health identifiers (matchable across data sets, but not reidentifiable by the evaluators), were made available to the evaluators by HVDHB. Qualitative feedback from interviews and focus groups further provided insight on benefits and/or liabilities of the solution, including influence on workflow and usability.

The eReferral use rose steadily to 1,000 transactions per month in 2008, thereafter showing moderate growth to 1,200 per month in 2010. Rate of eReferral from the community in 2010 is estimated at 56% of total referrals to the hospital from general practice, and as 71% of referrals from those having done at least one referral electronically. Figure 20.2 graphs the PIMS volumes for referral records indicating source as General Practice along with the transaction volume for all eReferrals (from the Clinical Workstation database) by year. A boost in total general practice referrals after relative stability in earlier years, tracking with increased eReferrals particularly between 2008 and 2009, indicates interaction of eReferral uptake and increase in total referrals.

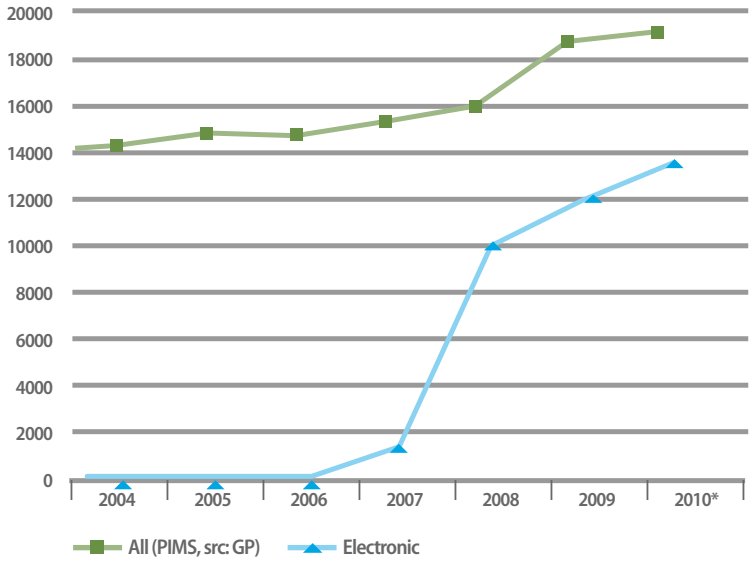


Figure 20.2. General practice referral volumes by year (* 2010 data inflated by 6/5ths to estimate full year volume).

Note. From "Introduction of electronic referral from community associated with more timely review by secondary services," by J. Warren, S. White, K. Day, Y. Gu, and M. Pollock, 2012, *Applied Clinical Informatics, 2*(4), p. 556. Copyright 2011 by Schattauer Publishing House. Reprinted with permission.

Referral latency from letter date to hospital triage improves significantly from 2007 to 2009 (Kolmogorov-Smirnov test, $p < 0.001$), from a paper referral median of eight days (inter-quartile range, IQR: 4–14) in 2007 to an eReferral median of five days (IQR: 2–9) and paper referral median of six days (IQR: 2–12) in 2009; see also Figure 20.3.

Qualitative feedback confirmed that the significant speed-up in referral processing shown in Figure 20.3 was achieved without changes in staffing levels. The evaluation concluded that substantial, rapid, and voluntary uptake of eReferrals was observed, associated with faster, more reliable, and more transparent referral processing. Clinical users appreciated improvement of referral visibility (status and content access); however, both GPs (referral senders) and specialists (receivers) point out system usability issues such as difficulties surrounding attachments in terms of both attaching at the sender's end and opening at the receiver's end.

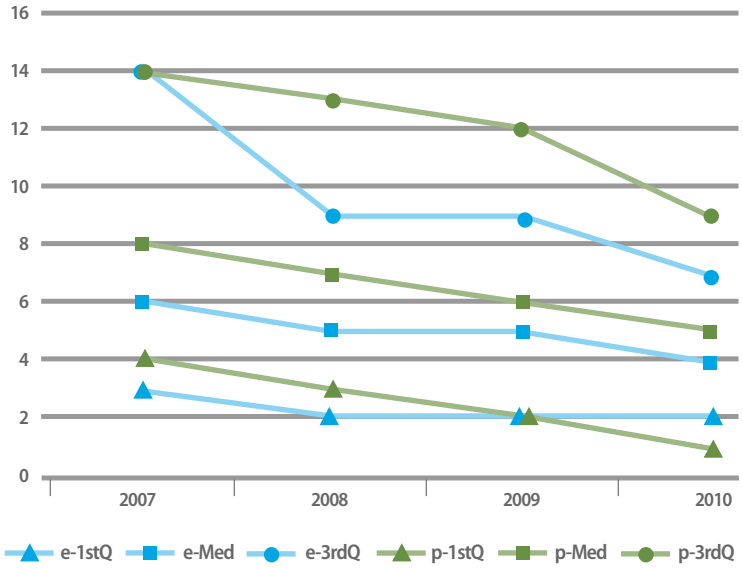


Figure 20.3. Median, first and third quartile ('Med', '1stQ' and '3rdQ' respectively) of letter-to-triage latency for e-referrals and paper referrals by year.

Note. From "Introduction of electronic referral from community associated with more timely review by secondary services," by J. Warren, S. White, K. Day, Y. Gu, and M. Pollock, 2012, *Applied Clinical Informatics, 2*(4), p. 557. Copyright 2011 by Schattauer Publishing House. Reprinted with permission.

20.2.2 Case Study Two — Promoting Sustained Use of a Shared Care Planning Program

The evaluation of New Zealand's National Shared Care Planning Program (NSCPP) was a case of Action Research-oriented evaluation that was planned during the eHealth program's business case stage in 2010. The evaluation was concurrently undertaken with the pilot development and implementation (2011 to 2012) with the aim to assess success as well as to support the pilot processes (Gu, Humphrey, Warren, & Streeter, 2014; Gu, Humphrey, Warren, Tibby, & Bycroft, 2012; Warren, Gu, & Humphrey, 2012). This example applied the principle that eHealth evaluation should begin before the new technology is introduced into the health workflow and be planned for along with the planning of the implementation itself. It demonstrated how evaluators could work in collaboration with the broader eHealth project team to understand and improve the user experience.

NSCPP took an IT-enabled approach to support shared care, shared decision-making, and care planning for long-term condition management. A Web-based technology solution was developed to provide a shared care record and coordination capability, including care plans, messages, and task assignment, for multidisciplinary care teams including patients themselves. The goal was to en-

able a patient-centred approach to care irrespective of the current care provider in general, specialist or allied healthcare settings, by facilitating both care coordination and supported self-management. The technology was integrated with GPs' PMS and has browser access to patient records for other community-based providers, hospital providers and patients.

NSCPP evaluation took an iterative action research approach applying both qualitative and quantitative methods. The pilot software was refined in response to ongoing feedback from the evaluation, which emphasizes attention to user feedback through interviews, focus groups and questionnaires to both participating healthcare professionals and patients, thematic analysis of communication records via the pilot system such as tasks and messages, as well as quantitative analysis of pilot system transaction records and health service usage data. Findings were used to identify any pressing issues and, according to evidence and expert experience, the corresponding recommendations for addressing the problems. This multifaceted data collection framework supported rapid synthesis of information and routine feedback loops to the program team to inform ongoing approaches in the rollout of the program. With an action research orientation, the methods and tools for the evaluation study were constantly examined and developed to accommodate the NSCPP development needs.

The program uptake, in terms of technology usage pattern and user experience, was closely monitored via qualitative feedback as well as analysis of user activities interacting with the pilot technology. And these were examined in the context of the users' professional roles, for example GP, general practice nurse, specialist physician, secondary nurse, allied health professional (including pharmacist and physiotherapist), and patient. Figure 20.4 captures user activities of creating and modifying tasks, notes, care plan elements and messages in the first nine months of the program, including the "Exploration" Phase, from March to June 2011 with one participating general practice and one secondary service and the "Limited Deployment" Phase (since July), extending to eight general practices, five secondary services and four community pharmacies. The modification activity includes marking a task as completed (which is the only such action available to patients in the patient portal at the time). Figure 20.5 shows user activities in terms of viewing the records, including tasks, notes, plans, messages, diagnosis, measurement results, medication and record summary, by month and roles.

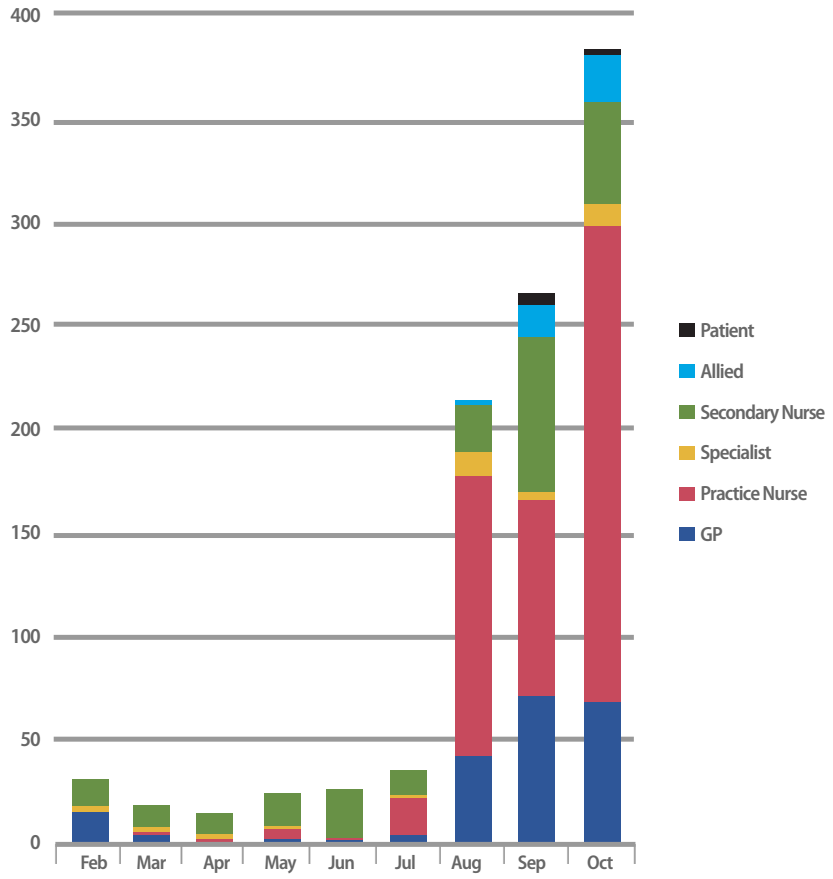


Figure 20.4. Sum of entries created or modified (over notes, care plan elements, messages and tasks) by role.

The above figures show the emergence of patient and allied users from August to October, with steady growth in allied health professional's activity and patient viewing. The role of specialist physicians as direct users, particularly with respect to element creation/modification, is quite small. The role of nurses is dominant for viewing and the creation or modification in all time periods except for a few cases in the early "Exploratory Phase" in February. In "Limited Deployment Phase" (since July), the role of general practice in element creation is highly dominant (at least two-thirds of entries), but is more balanced by other users with respect to viewing (roughly 50%). The observed pattern of nurses being the most active users extended to task assignment, an indicator of who is "driving". However, user interviews indicated that this might underestimate the guidance provided by physicians (indeed, at times, even literally looking over

the shoulder of the nurse operating the software). And, of course, the technology does not capture verbal communications that occur between nurses and physicians onsite. It is recognized that there is exciting potential for workforce transformation with NSCPP, but with the related challenge of defining the new responsibilities (and determining if these are met by new people or reorientation of existing roles). The question of who funds the time to create care plans had been raised repeatedly. A designated — and appropriately compensated — lead care coordinator (perhaps a nurse), would facilitate the solution of a further problem regarding the need to ensure timely responsiveness to issues emerging in the care of any given patient.

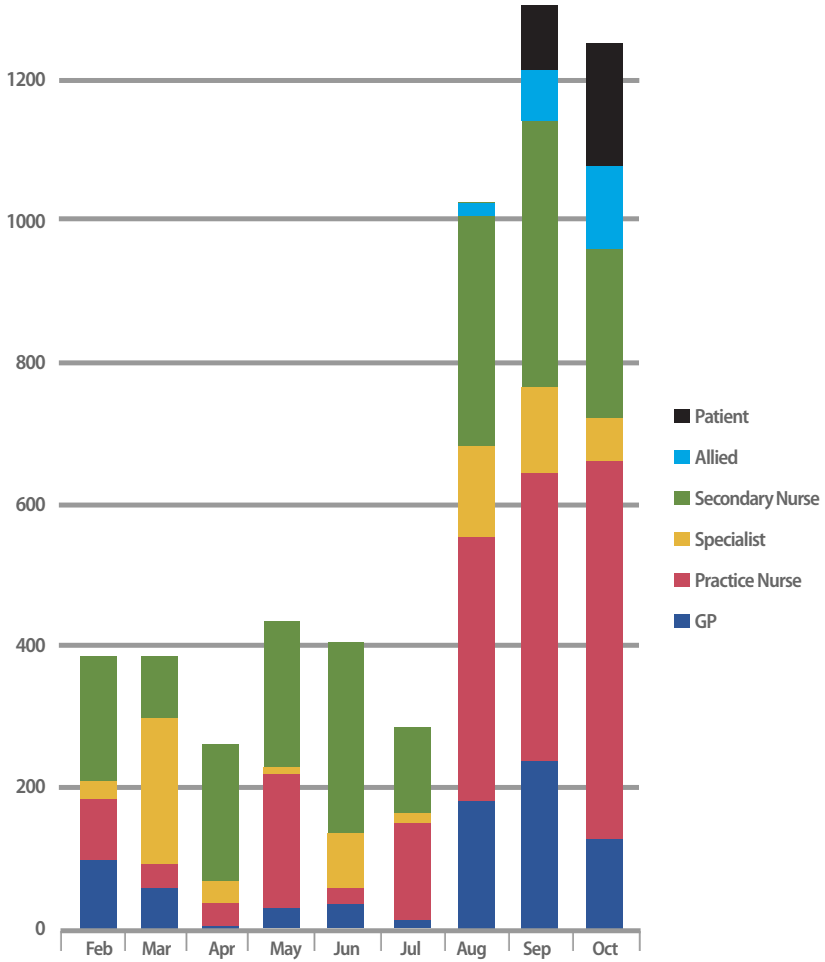


Figure 20.5. Elements viewed by user role based on number of NSCPP system audit log entries.

In fact, NSCPP has highlighted a range of fundamental challenges, including: (a) sociotechnical issues (e.g., interoperability in the non-standardized system environment, shortage of workforce skills to deliver care planning, lack of time/personnel to implement shared care, IT interface challenges, and mechanisms to involve patients and families); (b) governance of information, clinical workflows, privacy and funding models; and (c) patient safety concerns in relation to information access (and potentially input) by patients (e.g., detailed clinical communications can present difficulty for patient interpretation and are readily misunderstood). Moreover, there was no agreed definition consensus — in theoretic ground and among participating organizations or individuals — for "shared care planning" or the essential elements, roles or responsibilities needed for its delivery. On the other hand, most pilot participants acknowledged the notion that shared care is equally about sharing the care and responsibility with patients and their families as it is about sharing care within the interdisciplinary team. The NSCPP evaluation concluded that while many issues remain unresolved, the NSCPP experience is making the issues far more concrete and is building a wide community of clinical, and patient, users that now have first-hand experience to inform continued technology and policy development.

20.3 Discussion

Adoption is the essential first step in benefits realization for health IT. All evaluation studies must include investigation of adoption or risk misleading results. Substantial and sustained uptake in use of a system indicates success across a range of issues in project management, leadership, deployment, training, usability and overall "value proposition" of the system for the users. Conversely, failure in adoption indicates a breakdown. Continuing to pursue other aspects of evaluation in the face of poor adoption can lead to mistaking the IT system for the cause in a situation where other factors in fact account for observed variations in performance. Moreover, it is important to recognize that adoption is not an all-or-nothing proposition. Users may adopt some features of a system but not others, or uptake may be greater with one class of users than another (or at one site versus another). Such variation in uptake warrants more in-depth investigation and can lead to the discovery of opportunities for improvement wherever needed — in usability, training, system features or broader workflow and work role expectations.

Obtaining quantitative measures of adoption is usually relatively easy in the context of health IT because the systems, by their nature, lay down transactional "footprints" of their activity: a computerized physician order entry (CPOE) system creates records of orders, an electronic referral system creates records of referrals. Moreover, most systems will create usage logs for other purposes (e.g., security audit), although advanced planning to ensure the logging of the right information can greatly facilitate subsequent analysis.

The greatest benefits of adoption evaluation come, however, with the qualitative analysis in follow-up to areas where uptake is quantitatively weak. When interviewed, users can generally state the barriers to adoption that they are facing. Further, particularly if the interviews are structured to support such feedback, users may already have suggestions for improvements, or shed light on fundamental problems that underlie failure to adopt the system. This information can then be fed back to the broader project team to reduce the barriers as effectively as possible. As with all aspects of evaluation, the opportunities are greatest with a deployment process that is iterative and staged so that time and resources are available to learn from initial evaluation activities and to apply those lessons in later deployments.

20.4 Summary

In this chapter we have emphasized the importance of studying adoption in terms of substantial and sustained uptake of the system by its intended users as a cornerstone of eHealth evaluation. Through two case studies we have illustrated quantitative and qualitative approaches, with the quantitative dimensions underpinned by analysis of the quantity, source and timing of system transactions and the qualitative dimension underpinned by interviews. We have shown that uptake can be heterogeneous — with differences in usage levels for different classes of users, for example. We have seen that in follow-up with users about barriers to uptake we can achieve insights into both the current and potential role of the health IT-based innovation.

References

Chaudhry, B., Wang, J., Wu, S., Maglione, M., Mojica, W., Roth, E., . . . Shekelle, P. G. (2006). Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. *Annals of Internal Medicine, 144*(10), 742–752.

Eccles, M., McColl, E., Steen, N., Rousseau, N., Grimshaw, J., Parkin, D., & Purves, I. (2002). Effect of computerised evidence based guidelines on management of asthma and angina in adults in primary care: cluster randomised controlled trial. *British Medical Journal, 325*(7370), 941.

Finkelstein, J., Knight, A., Marinopoulos, S., Gibbons, M. C., Berger, Z., Aboumatar, H., . . . Bass, E. B. (2012, June). *Enabling patient-centered care through health information technology.* Evidence Reports / Technology Assessments No. 206. Rockville, MD: Agency for Healthcare Research & Quality.

Gu, Y., Humphrey, G., Warren, J., & Streeter, J. (2014). Measuring service utilisation impact of the Shared Care Planning Programme. *Health Care and Informatics Review Online, 18*(1), 16–24.

Gu, Y., Humphrey, G., Warren, J., Tibby, S., & Bycroft, J. (2012). *An innovative approach to shared care — New Zealand pilot study of a technology-enabled national Shared Care Planning Programme.* Paper presented at the Health Informatics New Zealand Conference, November 7 to 9, Rotorua.

Gu, Y., Warren, J., & Orr, M. (2014). The potentials and challenges of electronic referrals in transforming healthcare. *New Zealand Medical Journal, 127*(1398), 111–118.

Healthcare Information and Management Systems Society (HIMSS). (2015). Electronic Medical Record Adoption Model (EMRAM). Retrieved from http://www.himssanalytics.com/provider-solutions#block-himss-general-himss-prov-sol-emram

Kaplan, B., & Harris-Salamone, K. D. (2009). Health IT success and failure: recommendations from literature and an AMIA workshop. *Journal of the American Medical Informatics Association, 16*(3), 291–299. doi: 10.1197/jamia.M2997

Kitsiou, S., Manthou, V., Vlachopoulou, M., & Markos, A. (2010). Adoption and sophistication of clinical information systems in Greek public hospitals: Results from a national web-based survey. In P. Bamidis & N. Pallikarakis (Eds.), *XII Mediterranean Conference on Medical and Biological Engineering and Computing 2010* (Vol. 29, pp. 1011–1016). Berlin, Heidelberg: Springer.

Lau, F., Kuziemsky, C., Price, M., & Gardner, J. (2010). A review on systematic reviews of health information system studies. *Journal of the American Medical Informatics Association, 17*(6), 637–645. doi: 10.1136/jamia.2010.004838

Lester, H., & Campbell, S. (2010). Developing Quality and Outcomes Framework (QOF) indicators and the concept of 'QOFability'. *Quality in Primary Care, 18*(2), 103–109.

Marcotte, L., Seidman, J., Trudel, K., Berwick, D. M., Blumenthal, D., Mostashari, F., & Jain, S. H. (2012). Achieving meaningful use of health information technology: A guide for physicians to the EHR incentive programs. *Archives of Internal Medicine, 172*(9), 731–736. doi: 10.1001/archinternmed.2012.872

National Library of Medicine. (2014). 2014 MESH descriptor data: Referral and consultation. Retrieved from http://www.nlm.nih.gov/cgi/mesh/2014/MB_cgi?mode=&index=11500

Paré, G., & Sicotte, C. (2001). Information technology sophistication in health care: An instrument validation study among Canadian hospitals. *International Journal of Medical Informatics, 63*(3), 205–223. doi: 10.1016/S1386-5056(01)00178-2

Raymond, L., & Paré, G. (1992). Measurement of information technology sophistication in small manufacturing businesses. *Information Resources Management Journal, 5*(2), 4–16. doi: 10.4018/irmj.1992040101

Rousseau, N., McColl, E., Newton, J., Grimshaw, J., & Eccles, M. (2003). Practice based, longitudinal, qualitative interview study of computerised evidence based guidelines in primary care. *British Medical Journal, 326*(7384), 314.

Warren, J., Gu, Y., & Humphrey, G. (2012). *Usage analysis of a shared care planning system.* Paper presented at the The AMIA 2012 Annual Symposium (AMIA 2012), November 3 to 7, Chicago.

Warren, J., White, S., Day, K., Gu, Y., & Pollock, M. (2011). Introduction of electronic referral from community associated with more timely review by secondary services. *Applied Clinical Informatics, 2*(4), 546–564.

Chapter 21

Evaluation of Picture Archiving and Communications Systems

Don MacDonald, Reza Alaghehbandan, Doreen Neville

21.1 Introduction

Picture Archiving and Communication Systems (PACS) present an opportunity to radically change film-based radiology services both inside and outside the hospital setting. In the past, the usual medium for capturing, storing, retrieving and viewing radiology images was hard-copy film. The idea to replace film with digital images was first conceptualized in 1979 (Huang, 2003). However it was not until the early 1980s that advances in technology made introducing PACS into radiology departments feasible (Duerinckx, 2003). PACS replaces the film environment with an electronic means to seamlessly communicate and share radiology images and associated reports between health professionals.

21.2 Current State of Evidence

In evaluating the benefits of PACS there are many approaches and methodologies that can be employed. The approaches generally employed (i.e., formative versus summative or subjective versus objective) are basically different perspectives on how one can measure specific benefits. When reviewing the literature on PACS evaluations, it was rare that the author actually stated the approach taken in terms of evaluation perspectives. One must review the methodology closely to determine if, for example, the approach utilized was formative or summative. As well, in the majority of papers the methods section is limited to identifying the specific methods of data collection (e.g., surveys). In reviewing the literature, the methods most often used in evaluating PACS were: (a) questionnaires/surveys, (b) data collection sheets, (c) administrative data/project doc-

uments, (d) time and motion studies, (e) direct observation, (f) video recording, and (g) interviews. The environments in which PACS were most often evaluated were private clinics, radiology departments, and other hospital departments outside radiology.

21.2.1 Synthesis of Current Evidence

Many PACS evaluations published in the literature are not specific to a setting, but rather address a specific issue related to PACS. These include evaluations that investigated the following: financial benefits, pre-implementation planning, system integration, image quality, integration of voice recognition, and technical issues.

Financial benefits – Financial benefits that can be realized through the implementation of PACS fall into two areas, cost savings and increased revenues. In Canada, cost savings are achieved through the elimination (or reduction) of ongoing expenses related to the film environment, and are a direct result of the implementation of PACS. In the American health system if efficiencies are achieved with PACS over hard-copy film, additional revenues result for the given institution if the number of patients receiving radiology services is increased (i.e., increased patient throughput).

Pre-implementation planning – Planning for the implementation of PACS has drawn considerable interest from the research community in recent years. Pre-implementation planning studies have various degrees of scope, ranging from looking at the complete process, to carrying out a gap analysis and developing a Request for Proposals (RFP), to selecting the vendor (Ortiz & Luyckx, 2002; Swaton, 2002; Lepanto, Carrier, Gauvin, Dieumegarde, & Delage, 2002; Farnsworth, 2003; Bedel & Zdanowicz, 2004; Lawrence, 2005). Other implementation studies are even more specific, such as studies that investigate the role of a PACS Committee, the value of marketing PACS to end users (Viau, 2004), the challenge in linking PACS to external clinics (Arreola & Rill, 2003), and the degree of implementation of PACS in other countries (Foord, 2001; Inamura et al., 2001; Burbridge & Bell, 2004).

System integration – The maximum benefit of PACS is achieved when it is integrated into both the Hospital Information System (HOIS) and the Radiology Information System or RIS (Carrino et al., 1998; Reiner, Siegel, & Scanlon, 2002; Siegel & Reiner, 2003). A basic PACS architecture generally starts at the HOIS, as this is where patient demographic information is held and, in most cases, where the service order originates. Both patient demographic and order information is sent from the HOIS to the RIS, which distributes this information to the appropriate modality in the Radiology Department (e.g., chest X-ray). Once the image is created, it is sent from the RIS to the PACS for reviewing by the radiologist, who can then append the image report to the PACS (Mulvaney, 2002). The benefits of PACS integration into the RIS and HOIS systems include the elimination of redundant data entry, the availability of more accurate information in PACS, and a reduction in workload for radiology and clerical staff (Levine, Mun, Benson, & Horli, 2002).

Image quality – An increase in productivity and a reduction in costs are only beneficial if there is no loss of image quality when compared with traditional film. Given the massive amounts of computer memory (storage) required to store, transfer and retrieve digital images, earlier versions of PACS were disadvantaged simply because they were too expensive to operate (Agarwal, Rowberg, & Kim, 2003; Erickson, 2002). A relatively recent solution to the large amounts of space needed for digital imaging is to compress (or shrink) the image so that it does not require as much space for storage/transfer.

Integration of voice recognition – The installation of a voice recognition system that interfaces with PACS has been found to reduce the percentage of lost or unreported examinations (Hayt & Alexander, 2001) and improve report turnaround time (Azevedo-Marques, Carita, Benedicto, & Sanches, 2004). Voice recognition technology allows the radiologist to dictate an oral report via the voice recognition system, which is then attached to the appropriate image(s) in the PACS. The radiologist performs all the editing and corrections either by voice command or by manual typing (Marquez & Stewart, 2005). While voice recognition technology has made considerable advances in recent years, it still has some disadvantages. A particular concern is the potential for decreased face-to-face consultations between radiologists and physicians, given physicians have more immediate access to images and reports (Hayt & Alexander, 2001), and issues related to change management for both physicians and radiologists from multiple organizational perspectives.

Technical issues – Technical problems are always a possibility when new technology is introduced, and PACS is no exception. Problems with reliability of the system (Strickland, 2000) and delayed access to images (Inamura et al., 2001; Reed, Herzog, & Reed, 1996; Bryan, Weatherburn, Watkins, & Buxton, 1999) were identified in early studies of PACS. The issue of storage also garnered quite a bit of interest in the late 1980s and early 1990s, mainly because the digital image was so large and the storage capabilities were limited. Recent advances in technology have resolved the issue of storage (Naul & Sincleair, 2001), but other challenges still remain. These include access to historic images (Gamsu & Perez, 2003; Gaytos, Speziale, Bramson, & Treves, 2003), access to monitors and logging on to the system (Pilling, 2003), user friendliness (Cox & Dawe, 2002; Watkins, 1999; Krupinski, McNeill, Haber, & Ovitt, 2003), and overall IT support (Bedel & Zdanowicz, 2004; Hasley, 2002; Hayt, Alexander, Drakakis, & Berdebes, 2001).

21.2.2 Summary of Key Findings

The approaches to evaluating PACS are as diverse as the environments in which it is found. Areas that are usually studied focus in on increased efficiencies (e.g., increased report turnaround times — TAT) and productivity (e.g., more exams reported) and cost reduction. Cost is particularly important in the United States where a fee-for-service model is most prevalent.

21.3 Selected Case Study: Does PACS Improve Report Turnaround Times?

21.3.1 Introduction

In the fall of 2005, Canada Health Infoway and the Province of Newfoundland and Labrador (Canada) partnered on a $23 million initiative to implement one of the first province-wide Picture Archiving and Communication Systems in Canada. Prior to 2005, PACS implementations in Canada were funded either by provincial governments, regional health authorities, or such individual institutions as hospitals and clinics. In 1998, Newfoundland and Labrador initiated a project-based approach to implementing PACS, such that by 2005 approximately 70% of the service areas in the province had PACS capability. The challenge with the project-based approach was that these PACS were not interconnected and could not communicate beyond the local installation. To address these gaps, the 2005 PACS initiative in Newfoundland and Labrador was undertaken with two goals in mind: (a) to implement PACS in selected rural sites where no PACS currently existed, and (b) to address gaps in those regions where PACS was currently operational.

This section of the chapter describes a study specific to the impact that PACS had on turnaround times (TAT) for radiology reports. The Report TAT evaluation was carried out on the island portion of the province with a focus on hospitals in the two health authorities located on the east and west coasts.

21.3.2 Materials and Methods

This PACS Report TAT study was designed as a pre/post comparative benefits study. The majority of TAT data was collected from the hospitals' Radiology Information Systems (RIS), the Hospital Information System (HOIS), and the PACS: (a), each month, for minimum of three months pre-PACS implementation, and (b) each month, for minimum of nine months post-implementation, for a total of 12 data points. The mean TAT was derived for each pre/post period, excluding the month that PACS was implemented. A one-way analysis of variance (ANOVA) was used to determine if there was a statistically significant difference between the pre-PACS and post-PACS periods on the mean report TAT. The TAT was considered the dependent variable and pre/post time periods the independent variable. A p-value of < 0.05 would signify a significant difference in TAT between pre- and post-PACS.

21.3.3 Study Setting

The Province of Newfoundland and Labrador consists of two major geographical areas, the island of Newfoundland and a mainland section, Labrador. The province encompasses an area of 405,720 square kilometres with Labrador comprising 72% of the land area of the province, but containing only 5% of the population. The province is separated into four health authorities: Eastern Health Authority, Central Health Authority, Western Health Authority and the

Labrador/Grenfell Health Authority. The Eastern and Western Authorities were chosen for this study, given they had recently implemented PACS and pre/post exam data was available. The total population of Newfoundland and Labrador is approximately 525,000 (as of 2014), with the majority (300,000) residing in the Eastern Health Authority.

21.3.4 Report Total Turnaround Time

Defining an appropriate measure to study report TATs presented several challenges to the research team. Initially, we had hoped to measure the time the request for the exam was logged into the Radiology Information System (RIS), to the time the final report was posted back to the Hospital Information System (HOIS). However, several challenges became evident early into our study:

1. It was discovered that physicians sometimes utilized only the exams (or the draft reports) when providing patient care, thus minimizing the urgency of the radiologist to sign off on draft reports in a timely manner.

2. Some radiologists were known to verify all reports generated over an extended period of time on a particular day (e.g., every Friday afternoon).

3. Perhaps most importantly, check-in time was captured differently for inpatients and outpatients; that is, all inpatient "registrations" were recorded at 8:00 a. m. the morning after the physician had requested the exam. Conversely, outpatient "registrations" were recorded as the actual time the person registered in the hospital's radiology department.

Given the problems associated with our TAT measure, a modified measure was developed that excluded inpatient exams, and used the average monthly TAT for exams originating at outpatient registration to when the draft report was posted to the HOIS. At the time of the study, transcriptionists in many of the hospitals utilized a high-end tape recorder that was not interfaced with the HOIS. The transcriptionist reviewed the audiotape and typed the draft report directly into the HOIS. The radiologist then reviewed the draft report in the HOIS, made the necessary changes, and signed off on the report electronically.

Data for this modified TAT measure was collected for CT, echocardiography, MRI, nuclear medicine, general radiograph and ultrasound. In most cases, the collection period encompassed three (3) months pre-PACS implementation, and nine (9) months post-PACS implementation.

21.3.5 Results

Western Health Authority – Administrative data for all draft report TATs for outpatients was collected from the RIS and HOIS for each modality within scope in the Western Health Authority from September 2005 to December 2006 (N = 112,667). As a result of staggered implementation dates for PACS at the seven sites in the Western Health Authority, not all sites had complete data for three months pre- and nine months post-PACS implementation.

Eastern Health Authority – Administrative data for all draft report TATs for outpatients was collected from the RIS and HOIS for each modality within scope in the Eastern Health Authority for the period June 2004 to August 2005 (N = 177,855). As a result of staggered implementation dates for PACS at the three sites in the Eastern Health Authority, the pre-implementation and post-implementation periods differ depending on the month of implementation: June, July, or August 2004.

21.3.6 Discussion

The results of our study found that report TATs in some sites increased after PACS had been implemented, most notably in the Western Health Authority. In advance of discussing this anomalous finding, it is important to first consider PACS in the context of the enterprise of information systems that exist in today's modern hospitals. That is, there are a multitude of factors that need to be considered when investigating the benefits of PACS as it relates to report TATs. One needs to look at the entire enterprise, rather than PACS as a stand-alone system. Inamura and colleagues (1998) suggest the evaluation of PACS needs to look at the interaction between PACS, the Hospital Information System and the Radiology Information System, and how these systems interact with other information systems within the hospital. Foord (1999) concluded, "Installing PACS has very wide implications and it is important that these are well understood within the organisation and that acquiring a PACS is not seen as like buying another piece of imaging hardware, which has little functional impact on the radiology department and hospital as a whole" (p. 100). Reiner and Siegel (2002) identified several external factors to PACS which can impact on report TATs, such as facility type and size, HOIS/RIS/PACS integration, training, support staff, and patient population.

Another issue to be considered is what constitutes an acceptable TAT. The measure itself may be objective, however its interpretation is very subjective and can include many factors, such as the urgency of the event, the type of exam, hospital policy, staffing levels, exam volume and service environment (e.g., emergency department versus a chronic care unit). To put this into perspective, is a TAT of 100 hours for a non-urgent report any different than one of 50 hours? As one radiologist pointed out in follow-up to this issue, there is a big difference between statistical and clinical significance, and while there might be a statistically significant difference in an average TAT of 100 hours and one of 50 hours, as a physician treating a patient, the reduced time of 50 hours in the context of

100 hours may not be clinically significant if the case is non-urgent. The issue of clinical versus statistical significance is illustrated in a study carried out by Weatherburn, Bryan, Nicholas, and Cocks (2000) which found the rate of misdiagnosis pre-PACS was 1.5%, whereas the rate post-PACS was only 0.6%. The small difference raised this question: Regardless of whether the difference is statistically significant, is it clinically significant? The 1.5% rate of misdiagnosis suggests an efficiently run film environment existed in the emergency room prior to PACS being implemented. Following the implementation of PACS there was a statistical benefit realized, evident by the drop in misdiagnosis to 0.6%; however, this drop was not deemed to be clinically significant.

Western Health Authority – An analysis of the data obtained from the hospital information system at Hospital_A found that all six modalities under study experienced a significant increase in report TAT for the nine months following the implementation of PACS. This increase, as measured by the average TAT per month, was not entirely attributable to the initial high TATs for those months immediately following implementation. That is, it would be expected that longer TATs would be experienced immediately following the implementation of PACS given the inexperience of users. While there may be several reasons that contributed to the increased report TAT post-PACS at Hospital_A, an ongoing shortage of transcriptionists is believed to be the primary cause. At the time of the study, there was no voice recognition system at Hospital_A and all reports were recorded to a stand-alone recording system.

Of interest, many of the smaller peripheral sites in the Western Health Authority experienced decreases in report TATs following the implementation of PACS. Upon further investigation, it was determined that the most likely reason for this decrease was that before PACS was implemented, these sites would batch all their non-urgent exams (i.e., film) taken over a two- to three-day period, and then send them to Hospital_A via taxi for interpretation and reporting. Following the implementation of PACS, these exams were now available immediately to the radiologists at Hospital_A for reporting, thus eliminating the time previously taken in having the film transported over the road.

An important point to consider is that all sites within the Western Health Authority, with the exception of Hospital_A, have relatively small volumes of exams performed annually. To put this in context, the total exams within scope performed at the six peripheral sites in the Western Health Authority for the year under study was only 35,011, ranging from 1,134 to 16,727 per site. Adding in the volume of exams from Hospital_A (n = 77,656), the total volume of exams for the Western Health Authority was only 112,667.

Eastern Health Authority – In the Eastern Health Authority there were three hospitals for which TAT data was collected pre- and post-PACS implementation. Hospital_H carried out 97,922 exams for those modalities within scope, Hospital_I 73,428, and Hospital_J 6,505.

Hospital_H provided report TAT data pre- and post-PACS for CT, echocardiography, MRI, nuclear medicine, general radiograph, and ultrasound. All modal-

ities, with the exception of nuclear medicine, experienced a reduction in average TAT for the three months pre-PACS implementation compared to the 12 months post-PACS. Similar to Hospital_A in the Western Authority, Hospital_H also experienced issues related to a lack of transcriptionists. However, given the larger size of Hospital_H compared to Hospital_A, the impact of a reduction in transcriptionists was partially mitigated by the fact Hospital_H had more transcriptionists on staff to share the workload. In addition, the administration at Hospital_H introduced short-term measures to address the delay in TATs, including increasing overtime and contracting with retired transcriptionists.

Hospital_I exams within scope included CT, echocardiography, nuclear medicine, general radiograph, and ultrasound. Only TATs for nuclear medicine and general radiographs experienced a decrease from pre- to post-PACS, whereas the average TAT for the other three modalities remained statistically the same. In investigating why some modalities experienced a decrease in TAT, while others apparently did not, no one cause was identified. The problem the research team experienced in carrying out such investigations was that administrative databases are limited when one wants to study cause and effect. However, one explanation put forward by one Director of Radiology was a likely reduction in human resources (i.e., radiologists and transcriptionists) available, either through retention or illness, for extended periods of time for the year that TAT data was collected. During these times of staff shortages, it is possible that the reporting of some types of exams were given priority over others. Another reason may be specific hospital policies that dictate which exams are given priority for reporting

Hospital_J is a psychiatric hospital that also provides general radiographs to the general public through a pre-appointment outpatient setting. Over the study period, there were 6,505 general radiology exams performed at this site, with a decrease in report TAT found from pre- to post-PACS. Hospital_J has two technologists on staff, and no radiologist. In the film environment, a radiologist would visit Hospital_J twice a week to report on all exams taken since the previous visit. In the PACS environment, the technologists now only need to call a radiologist at one of the other sites and let them know that the exam is now posted on PACS and request a consult. The ability to post exams on PACS for external review was the most significant factor in reducing report TATs at Hospital_J.

21.3.7 Conclusion

The implementation of PACS in two of the four health regions in Newfoundland and Labrador had mixed results with respect to Report TATs. Our study found that increases in report TATs in some smaller sites following the implementation of PACS was due mainly to a lack of support staff (transcriptionists), rather than the PACS itself. In the larger sites studied, a lack of transcriptionists was also evident; however, the impact on TATs was less profound given the reductions in support staff in the Radiology Department could be mitigated through other resources available in the larger sites.

21.4 Issues, Guidance and Implications

Where sites are paid for each radiology service provided to a patient (e.g., the United States), PACS can provide an opportunity to increase revenues. This is made possible when radiologists become more efficient in reviewing digital exams (images) and preparing reports for referring physicians. With this increased efficiency, hospitals can accommodate more new patients (i.e., increase productivity) from their pool of referring physicians (Reed et al., 1996; Chopra, 2000; Kim, Park, Chun, & Nam, 2002; Andriole, Rowberg, & Gould, 2002; Hunt, 1998). In Canada, the delivery of health services is funded through the Canada Health Transfer (CHT), which provides universal health care insurance to all residents of Canada. Therefore, PACS provides limited opportunity for hospitals in Canada to generate revenues by increasing the number of patients seeking radiology services. Nevertheless, from an accountability perspective, investments in health information systems are costly and it is necessary to quantify the success of such systems and the degree to which the investment was justified (Protti, 2002). Challenges to addressing these concerns include:

1 Efficiency (doing things right) is easier to measure than effectiveness (doing the right thing).

2 New systems are intended to change difficult-to measure actions.

3 Strategic systems elude measurement.

4 Infrastructure investments are difficult to justify on a return on investment (ROI) basis.

Adding to the challenge is that the literature is not conclusive on whether PACS can actually result in savings and/or increase revenues and profits (Strickland, 2000; Maass, Kosonen, & Kormano, 2001; Maass et al., 2002; Grosskopf, 1998; Terae, Miyasaka, Fujita, & Shirato, 1998; Cartier, 1999; Andriole et al., 2002; Colin et al., 1998; Nitrosi et al., 2007), given that the level of benefit achievement depends on a multitude of confounding factors, such as the funding model in place, the degree of HOIS/RIS/PACS integration, the level of training and support staff, the size and type of the PACS site, and the population served (Reiner, Siegel, Carrino, & Goldburgh, 2002), and how efficient the film site was before PACS was implemented (Lepanto et al., 2002).

The volume of exams performed in a site, and its relationship to the expected benefits of PACS, can also impact on the level of benefits achieved through introducing PACS. While installing PACS in a site that only averages 10,000 exams per year may not be a practical investment for most sites, it nevertheless raises the question of what constitutes the necessary volume of images before an investment in PACS becomes feasible. An earlier study by Bauman, Gell, and Dwyer (1996) stated that a large PACS installation required a minimum of

20,000 examinations per year to ensure the feasibility of PACS, whereas seven years later Siegel and Reiner (2003) reported the cut-off was at 39,000 exams. In classifying sites, Cartier (1999) carried out a study in a "small" hospital that produced 15,000 exams a year, while Hayt et al. (2001) carried out a study in a "large" hospital that produced 116,000 exams per year. While these studies classified the size of a site either in relation to the number of beds, or the actual volume of exams, there is no consensus on standards for such classifications. Nevertheless, such studies do raise the question of how to interpret the benefits of PACS within the context of exam volume.

Capital and operational factors associated with the implementation of PACS in the Western Health Authority were very costly. The most significant contributors to the cost of PACS, and the main reason for not realizing a financial return on investment, were equipment and maintenance costs. In the Western Health Authority total cost of PACS was $4.1 million, of which $2.4 million was for hardware (58%). In addition to hardware costs, annual licensing and maintenance costs usually run about 10% to 15% of capital costs, which in the case of the Western Health Authority came to $229,000 per year. One potential opportunity to reduce PACS equipment costs is for multiple sites to partner and offer a joint request for proposals (RFP), thus taking advantage of any economies of scale. The overall cost for the provincial implementation and/or enhancement of PACS was $23 million, not an insignificant amount, even nationally. Yet even with this expenditure, there were no major savings realized, and the costs of the PACS equipment resulted in most hospitals in the province not achieving a return on investment. Until PACS hardware, software and licensing fees comes down in price, it is unlikely, except in the largest urban hospitals, that there will be any financial return on investment for the majority of PACS implemented in Canada.

21.5 Summary of Evaluation Issues

The real challenge is not in determining revenues and/or savings, although both are important and relatively easy to measure. The challenge is determining the indirect benefits of PACS that even today continue to elude meaningful measurement. That is, how can one quantify in financial terms benefits such as improved patient care or outcomes, improved access, or clinician satisfaction? In spite of the 25-plus years of PACS research, there still is no consistent evidence that supports the financial benefits across the many diverse environments in which PACS operates. Sites having high exam volumes, inefficient film environments, and opportunities to generate revenues, offer the best likelihood of achieving a financial return on investment. In contrast, the Western Health Authority had a moderate exam volume, an efficiently run film environment, and no opportunities for generating revenue. This environment resulted in the cost per case analysis in Western Health Authority concluding that unless the planning horizon is lengthy, PACS is more expensive to operate than within the traditional film environment.

21.5.1 Guidance for Future Directions

While this study focused on a report TATs for PACS, it is recognized that the true benefits of PACS are quite far-reaching. There are many other benefits of PACS that need to be considered in the broader context of patient care. Improved efficiency and productivity, which are achieved in part through improved report turnaround times (Azevedo-Marques et al., 2004; Reiner & Siegel, 2002; Mackinnon, Billington, Adam, Dundas, & Patel, 2008) and immediate access to reports and images from multiple sites 24 hours a day, seven days a week (Watkins, 1999; Mackinnon et al., 2008; Bryan et al., 1998; Ravin, 1990; Srinivasan, Liederman, Baluyot, & Jacoby, 2006; Hurlen, Ostbye, Borthne, & Gulbrandsen, 2010; Bolan, Guimaraes, & Mueller, 2008) are but two benefits of PACS considered to offset any higher costs for PACS.

21.5.2 Policy and Practice Implications

From a clinical practice perspective, many of the aforementioned benefits feed into the decision of a hospital/clinic to budget for the changeover from film to PACS, as such costs generally come for the hospital/clinic's operating budget. Policy really does not have as big an influence on moving to PACS at the institutional level as that of clinical benefits. One area that would have broader interest beyond the pure clinical piece is when PACS is considered a valuable tool for recruiting and retaining hard-to-find radiologists in a very competitive national and international market. It is understandable that a radiologist looking for employment will likely go to an environment where the latest technology is available (and stay there), and PACS certainly delivers in that sense.

21.6 Summary

PACS has been available for more than a quarter of a century yet it is still difficult, if not impossible, to measure its true benefits, given the differences in evaluation approaches and clinical environments. It is also difficult to separate PACS from all the other information systems that operate within any environment. Perhaps even more difficult is to attempt to evaluate PACS (as with any health information system) from a financial perspective, given the difficulty in quantifying and defining a price on improved quality of care for our patient population. Most PACS evaluations examine improved efficiencies or productivity, and these then become proxies for improved quality of care and, ultimately, improved health outcomes. For now, we will have to take that assumption on faith.

References

Agarwal, A., Rowberg, A. H., & Kim, Y. (2003). Fast JPEG 2000 decoder and its use in medical imaging. *IEEE Transactions on Information Technology in Biomedicine, 7*(3), 184–190.

Andriole, K. P., Rowberg, A. H., & Gould, R. G. (2002). Workflow assessment of digital versus computed radiography and screen-film in the outpatient environment. *Journal of Digital Imaging, 15*(1), 124–126.

Arreola, M., & Rill, L. N. (2003). Enterprise-wide CR implementation: The Shands Healthcare System experience. *Journal of Digital Imaging, 16*(2), 173–179. doi: 10.1007/s10278-003-1653-z

Azevedo-Marques, P. M., Carita, E. C., Benedicto, A. A., & Sanches, P. R. (2004). Integrating RIS/PACS: the web-based solution at University Hospital of Ribeirao Preto, Brazil. *Journal of Digital Imaging, 17*(3), 226–233.

Bauman, R., Gell, G., & Dwyer, S. (1996). Large picture archiving and communication systems of the world — Part I. *Journal of Digital Imaging, 9*(3), 99–103.

Bedel, V., & Zdanowicz, M. (2004). PACS strategy for imaging centers. *Radiology Management, 26*(5), 24–29.

Bolan, G. W., Guimaraes, A. S., & Mueller, P. R. (2008). Radiology report turnaround: Expectations and solutions. *European Society of Radiology, 18*(7), 1326–1328.

Bryan, S., Weatherburn, G., Watkins, J., & Buxton, M. (1999). The benefits of hospital-wide picture archiving and communication systems: a survey of clinical users of radiology services. *The British Journal of Radiology, 72*(857), 469–478.

Bryan, S., Weatherburn, G., Watkins, J., Roddie, M., Keen, J., Muris, N., & Buxton, M. (1998). Radiology report times: Impact of picture archiving and communication systems. *American Journal of Roentology, 170*(5), 1153–1159.

Burbridge, B., & Bell, C. (2004). The digital readiness of imaging facilities in Saskatchewan. *Canadian Association of Radiologists, 55*(5), 311–314.

Carrino, J. A., Unkel, P. J., Miller, I. D., Bowser, C. L., Freckleton, M. W., & Johnson, T. G. (1998). Large-scale PACS implementation. *Journal of Digital Imaging, 11*(3), 3–7.

Cartier, M. (1999). The challenge of PACS in a small hospital. *Radiology Management, 21*(4), 23–28.

Chopra, R. M. (2000). Why PACS is no longer a four-letter word. *Radiology Management, 22*(5), 44–48.

Colin, C., Vergnon, P., Guibaud, L., Borson, O., Pinaudeau, D., Perret Du Cray, M. H., Pasquier, J. M., & Tan-Minh, V. (1998). Comparative assessment of digital and analog radiography: diagnostic accuracy, cost analysis and quality of care. *European Journal of Radiology, 26*(3), 226–234.

Cox, B., & Dawe, N. (2002). Evaluation of the impact of a PACS system on an intensive care unit. *Journal of Management in Medicine, 16*(2/3), 199–205.

Duerinckx, A. (Ed.). (2003, March). Introduction to two PACS '82 panel discussions: "Equipment manufactures' view on PACS" and "The medical community's view on PACS". *Journal of Digital Imaging, 16*(1), 29–31.

Erickson, B. (2002). Irreversible compression of medical images. *Journal of Digital Imaging, 15*(1), 5–14.

Farnsworth, T. J. (2003). PACS for imaging centers. *Radiology Management, 25*(3), 36–41.

Foord, K. (1999). PACS: the second time around. *European Journal of Radiology, 32*(2), 96–100.

Foord, K. (2001). Year 2000: status of picture archiving and digital imaging in European hospitals. *European Journal of Radiology, 11*(3), 513–524.

Gamsu, G., & Perez, E. (2003). Picture archiving and communication systems (PACS). *Journal of Thoracic Imaging, 18*(3), 165–168.

Gaytos, D., Speziale, J., Bramson, R., & Treves, S. T. (2003). The PACS pre-implementation process at a major teaching hospital: A multi-disciplinary approach. *Journal of Digital Imaging, 16*(Suppl 1), 73–80.

Grosskopf, E. J. (1998). On the road to PACS: Confronting the issues. *Radiology Management, 20*(5), 26–33.

Hasley, T. (2002). PACS support: The radiology approach. *Radiology Management, 24*(6), 26–30.

Hayt, D. B., & Alexander, S. (2001). The pros and cons of implementing PACS and speech recognition systems. *Journal of Digital Imaging, 14*(3), 149–157.

Hayt, D. B., Alexander, S., Drakakis, J., & Berdebes, N. (2001). Filmless in 60 days: the impact of picture archiving and communications systems within a large urban hospital. *Journal of Digital Imaging, 14*(2), 62–71.

Huang, H. K. (2003). Some historical remarks on picture archiving and communication systems. *Computerized Medical Imaging and Graphics, 27*(2/3), 93–99.

Hunt, D. (1998). Making PACS the present, not the future. *Radiology Management, 20*(5), 35–38.

Hurlen, P., Ostbye, T., Borthne, A., & Gulbrandsen, P. (2010). Introducing PACS in the late majority: A longitudinal study. *Journal of Digital Imaging, 23*(1), 83–94. doi: 10.1007/s10278-008-9160-x

Inamura, K., Umeda, T., Sukenobu, Y., Matsuki, T., Kondo, H., Takeda, H., … Kozuka, T. (1998). HIS/RIS contribution to image diagnosis and maximization of efficacy of PACS when coupled with HIS/RIS. *Computer Methods and Programs in Biomedicine, 57*(1-2), 41–49.

Inamura, K., Konishi, J., Nishitani, H., Kousaka, S., Matsumura, Y., Takeda, H., & Kondoh, H. (2001). Status of PACS and technology assessment in Japan. *Computer Methods and Programs in Biomedicine, 66*(1), 5–15.

Kim, S. A., Park, W. S., Chun, T. J., & Nam, C. M. (2002). Association of the implementation of PACS with hospital revenue. *Journal of Digital Imaging, 15*(4), 247–253.

Krupinski, E., McNeill, K., Haber, K., & Ovitt, T. (2003). High-volume teleradiology service: focus on radiologist satisfaction. *Journal of Digital Imaging, 16*(2), 203–209.

Lawrence, P. (2005). Preparing a cultural strategy for PACS. *Radiology Management, 27*(1), 21–26.

Lepanto, L., Carrier, R., Gauvin, A., Dieumegarde, M., & Delage, M. (2002). Planning an enterprise-wide PACS: the Centre Hospitalier de l'Université de Montréal experience. *Journal of Digital Imaging, 15*(Suppl 1), 252–253.

Levine, B. A., Mun, S. K., Benson, H. R., & Horli, S. C. (2003). Assessment of the integration of a HIS/RIS with a PACS. *Journal of Digital Imaging, 16*(1), 133–140.

Maass, M., Kosonen, M., & Kormano, M. (2001). Cost analysis of Turku University Central Hospital PACS in 1998. *Computer Methods and Programs in Biomedicine, 66*, 41–45.

Maass, M., Kosonen, M., Kormano, M., Chan, L., Trambert, M., Kywi, A., & Hartzman, S. (2002). PACS in private practice — effect on profits productivity. *Journal of Digital Imaging, 15*(Suppl 1), 131–136.

Mackinnon, A. D., Billington, R. A., Adam, E. J., Dundas, D. D., & Patel, U. (2008). Picture archiving and communication systems lead to sustained improvements in reporting times and productivity: results of a 5-year audit. *Clinical Radiology, 63*(7), 796–804.

Marquez, L. O., & Stewart, H. (2005). Improving medical imaging report turnaround times: The role of technology. *Radiology Management, 27*(2), 26–31.

Mulvaney, J. (2002). The case of RIS/PACS integration. *Radiology Management, 24*(3), 24–29.

Naul, L. G., & Sincleair, S. T. (2001). Radiology goes filmless. *Postgraduate Medicine, 109*(6), 107–110.

Nitrosi, A., Borasi, G., Nicoli, F., Modigliani, G., Botti, A., Bertolini, M., & Notari, P. A. (2007). Filmless radiology department in a full digital regional hospital: Quantitative evaluation of the increased quality and efficiency. *Journal of Digital Imaging, 20*(2), 140–148.

Ortiz, A. O., & Luyckx, M. P. (2002). Preparing a business justification for going electronic. *Radiology Management, 24*(1), 14–21.

Pilling, J. R. (2003). Picture archiving and communication systems: the users' view. *The British Journal of Radiology, 76*(908), 519–524.

Protti, D. (2002). What can the American electronic health record (EHR) pioneers tell us about what it takes to be successful? *Healthcare Management Forum, 15*(2), 33–35.

Ravin, C. E. (1990). Initial experience with automatic image transmission to an intensive care unit using picture archiving and communications system technology. *Journal of Digital Imaging, 3*(3), 195–199.

Reed, D. H., Herzog, D. G., & Reed, G. (1996). A visual guide to PACS. *Radiology Management, 18*(4), 20–22.

Reiner, B., Siegel, E., Carrino, J., & Goldburgh, M. (2002). SCAR radiologic technologist survey: Analysis of the impact of digital technologies on productivity. *Journal of Digital Imaging, 15*(3), 132–140.

Reiner, B., Siegel, E., & Scanlon, M. (2002). Changes in technologist productivity with implementation of an enterprise wide PACS. *Journal of Digital Imaging, 15*(1), 22–26.

Reiner, B. I., & Siegel, E. L. (2002). Technologists' productivity when using PACS: comparison of film-based versus filmless radiography. *American Journal of Roentology, 179*(1), 33–37

Siegel, E. L. & Reiner, B. I. (2003). Filmless radiology at the Baltimore VA Medical Center: a 9-year retrospective. *Computerized Medical Imaging and Graphics, 27*(2), 101–109.

Srinivasan, M., Liederman, E., Baluyot, N., & Jacoby, R. (2006). Saving time, improving satisfaction: the impact of a digital radiology system on physician workflow and system efficiency. *Journal of Healthcare Information Management, 20*(2), 123–132.

Strickland, N. H. (2000). PACS (picture archiving and communication systems): filmless radiology. *Archives of Disease in Childhood, 83*(1), 82–86.

Swaton, N. (2002). Learn from experience: insight of 200+ PACS customers. *Radiology Management, 24*(1), 22–27.

Terae, S., Miyasaka, K., Fujita, N., & Shirato, H. (1998). A hospital-wide PACS: 7 year experience and new development. *Computer Methods and Programs in Biomedicine, 57*(1-2), 5–12.

Viau, M. A. (2004). PACS — and beyond. *Radiology Management, 26*(3), 36–40.

Watkins, J. (1999). A hospital-wide picture archiving and communication system (PACS): the views of users and providers of the radiology service at Hammersmith Hospital. *European Journal of Radiology, 32*(2), 106–112.

Weatherburn, G., Bryan, S., Nicholas, A., & Cocks, R. (2000). The effect of a picture archiving and communications system (PACS) on diagnostic performance in the accident and emergency department. *Journal of Accident and Emergency Medicine, 17*(3), 180–184.

Chapter 22

Evaluation of Provincial Pharmacy Network

Don MacDonald, Khokan C. Sikdar, Jeffrey Dowden, Reza Alaghehbandan, Peizhong Peter Wang, Veeresh Gadag

22.1 Introduction

Adverse drug events (ADEs) are a concern in both inpatient (Evans, Lloyd, Stoddard, Nebeker, & Samore, 2005; Bates et al., 1995; Baker et al., 2004) and outpatient (Budnitz et al., 2005; Zed et al., 2008) settings. An ADE is defined as an iatrogenic hazard or incident that is created either through omission or commission of the administration of a drug or drugs (prescription or non-prescription), harming a patient whose outcome is always unexpected and unacceptable to the patient and healthcare provider (Tafreshi, Melby, Kaback, & Nord, 1999; Nebeker, Hoffman, Weir, Bennett, & Hurdle, 2005). Such events are a significant cause of morbidity and mortality (Juntti-Patinen & Neuvonen, 2002; Alexopoulou et al., 2008), and result in significant resource utilization, including increased emergency room (ER) and physician visits, diagnostic tests, medication use, and hospital admissions. Studies conducted in the United States estimate that such events account for 17 million ER visits and 8.7 million hospital admissions each year (Bates et al., 1997; Johnson & Bootman, 1995). Between 1995 and 2000, costs associated with ADEs rose from US$76.6 billion to over US$177.4 billion (Johnson & Bootman, 1995; Ernst & Grizzle, 2001). It would be expected that if a Pharmacy Network were deployed and that it included complete medication profiles and automatic drug utilization reviews, then adverse drug events resulting in an ER visit would be reduced in the population.

22.2 Current State of Evidence

ADEs have been mostly studied among patients admitted to hospital, and it has been estimated that 5% to 25% of hospital admissions are drug-related (Samoy et al., 2006; Pirmohamed et al., 2004). However, ADEs occurring in outpatient settings and treated in ERs receive less attention, even though more than 80% of community-dwelling adults use medications on a weekly basis, and approximately threefold more patients are treated in ERs for ADEs compared to those admitted to hospital (Budnitz et al., 2005; McCaig & Burt, 2003; Kaufman, Kelly, Rosenberg, Anderson, & Mitchell, 2002). The Institute of Medicine (1999) report, *To Err Is Human: Building a Safer Health System*, concluded that the solution to preventing medical errors is "building a safer health system" that identifies patient safety as a prerequisite to high-quality care. Despite widespread recognition of the need for a safer health system, ADEs occurring in community settings remain a substantial cause of ER visits. A Pharmacy Network is a regional drug information system that offers population-based online, real-time medication profiles and an interactive database to assist pharmacists and physicians in producing optimal medication treatment. Such networks would also provide the tools to monitor, track, and mitigate ADEs and medication errors that occur in the community.

22.2.1 Synthesis of Current Evidence

ADEs are a major public health problem given that such events are the most common type of injuries experienced by hospitalized patients (Institute of Medicine, 1999). ADEs may lead to hospitalization, or occur during hospitalization and contribute to an increased length of stay. The recent focus on patient safety and the concern about the number of negative outcomes resulting from drug use, rather than the underlying diseases, has prompted health care professionals to take a critical look at these drug responses.

A series of studies examined ADEs among hospitalized patients in the United States and Australia (Bates et al., 1995; Lazarou, Pomeranz, & Corey, 1998; Roughead, Gilbert, Primrose, & Sansom, 1998; McDonnell & Jacobs, 2002; Zhang et al., 2009); however, less research is available about these events in hospitalized patients in Canada. A U.S.-based meta-analysis by Lazarou et al. (1998) revealed that the incidence of serious ADEs in hospitalized patients was 2.1%, while for those newly admitted to a hospital it was 4.7%. A subsequent study reported ADEs were between the fourth and sixth leading cause of death (Tafreshi et al., 1999). Other studies have found ADEs occurred in between 2% and 20% of hospitalized patients (Roughead et al., 1998; McDonnell & Jacobs, 2002; Zhang et al., 2009). Baker and colleagues (2004) provided a national estimate of the incidence of adverse events among adult patients in Canada (7.5 per 100 hospital admissions). After extrapolating to the entire population of Canada, the number of hospital admissions attributed to adverse events was estimated between 141,250 and 232,250 in 2000 (Baker et al., 2004). Further-

more, Canadian incident reporting data indicated a 35% increase of adverse reactions from 2008 to 2009 (Health Canada, 2010).

ADEs are common and can have serious consequences in an older population. According to recent population estimates, Canadians 65+ population grew by 12% between 2001 and 2006 and this demographic now represents about 16% of the total population (Statistics Canada, 2007; Canadian Institute for Health Information, 2010; CBC News, 2015). Elderly individuals are vulnerable to ADEs because of their multiple drug consumption patterns and biologic changes, which may restrict their drug consumption and inhibit physiological processes they take to manage multiple comorbid conditions and because of pharmacokinetics and pharmacodynamics changes (Zhang et al., 2009; Bates, 1998). Furthermore, ADEs can be recurrent events, in that an individual may experience one or more such events over a period of time. It is important to identify the magnitude of ADEs in this high-risk group to aid physicians in their decisions about prescribing, delivering, administering, and monitoring drug therapies. If predictive factors can be identified, this would allow providers to identify early symptoms of ADEs and to offer rapid response to the patient (Field et al., 2001).

Although prior research (Zhang et al., 2009; Field et al., 2001; French, 1996; Fialová et al., 2005; Onder et al., 2002; Onder et al., 2003; Chrischilles, VanGilder, Wright, Kelly, & Wallace, 2009) has identified several risk factors for the occurrence of ADEs among older adults (e.g., age, sex, and drug regimen), little is known about the risk factors associated with recurrent ADEs. For public health planning and the evaluation of quality management programs, it is important to study recurrent ADEs, rather than only the first event (Donaldson, Sobolev, Cook, Janssen, & Khan, 2009). Given the risk of both health service utilization and the patient's burden of illness increasing with each subsequent ADE, the number of ADEs is a more robust indicator of risk than a single event (Glynn & Buring, 1996).

22.2.2 Summary of Key Findings

ADEs have been mostly studied among patients admitted to hospital, and it has been estimated that between 5% and 25% of hospital admissions are drug-related. However, ADEs occurring in outpatient settings and treated in ERs receive less attention, even though more than 80% of community-dwelling adults use medications on a weekly basis, and approximately threefold more patients are treated in ERs for ADEs compared to those admitted to hospital.

22.3 Selected Case Study – Adverse Drug Events in Adult Patients Leading to ER Visits

22.3.1 Setting and Study Population

The study setting was two adult acute care hospitals, the Health Science Centre (HSC) and St. Clare's Mercy Hospital (SCMH), both of which deliver tertiary care

in the capital city of St. John's in Newfoundland and Labrador (N.L.), Canada. These two hospitals serve a catchment area of approximately 280,000 residents, and together have an average of 28,000 acute separations and 80,000 ER visits per year. Both hospitals capture electronic summary data on all ER visits in an emergency room triage database. Eligible subjects for this study included all patients aged 18 years or over that were residents of N.L. and presented to one of the two ERs between January 1, 2005 and December 31, 2005.

ER visits with a high probability of not being due to an ADE (e.g., motor vehicle accident, substance abuse, drug abuse, attempted suicide, cut- or burn-related injuries, etc.) were excluded. It should be noted that conditions such as attempted suicide and drug abuse would not likely be the presenting complaint. Therefore, these ER visits may not have been excluded from the sampling frame; rather they were excluded later (if selected) from the study sample during the chart review phase of the study. Patients who presented to ERs through a referral process, but were subsequently identified as a valid ER visit, were included in this study.

22.3.2 Study Sample

Charts were selected from the sampling frame using a stratified random sampling design. There were six strata based on patients' sex and age at ER visit (Male 18 to 44, Male 45 to 64, Male 65+, Female 18 to 44, Female 45 to 64, and Female 65+). Evidence in the literature regarding the prevalence of ADEs in admissions was found to be inconsistent, ranging from 5% to 25%, which can be mostly attributed to differences in study designs and patient demographics (Kaufman et al., 2002; Institute of Medicine, 1999; Lazarou et al., 1998). We estimated 10% of ER visits would be attributed to ADEs in patients aged 18 years and older. To achieve a 95% confidence interval (±4%), we determined that we would need a sample size of 217 ER visits for each stratum, resulting in a total of sample 1,302 ER visits. To reduce the sampling error, and to compensate for the exclusion of ER visits that would be attributed to suicide attempts and drug abuse, we added a 10% over-sample to the sample. After the chart review was completed, the final sample size for the study was 1,458, resulting in a 12% oversample. This difference of 2% was attributed to inclusion of ER visits through referrals. For patients with multiple ER visits during the study period, only one visit was selected at random as the index visit for review.

22.3.3 Outcomes and Definitions

An ADE is defined as any undesirable effect caused by the interaction of a drug (prescription or non-prescription) with a patient (Morimoto, Gandhi, Seger, Hsieh, & Bates, 2004). Events may be the result of normal or inappropriate use of a medication, and could range from minor reactions such as a skin rash to serious and life-threatening events, even death. Medication errors (MEs) are mishaps that occur during prescribing, transcribing, dispensing, administering, adherence, or monitoring a drug. Medication errors are more common than

adverse drug events, but result in harm less than 1% of the time, with about 25% of adverse drug events attributed to medication errors (Nebeker, Barach, & Samore, 2004). We studied ADEs, defined as "injury resulting from the use of a drug" (Nebeker et al., 2005) that encompasses all traditional adverse effects plus harm from any MEs. We also used "possible adverse drug event" (PADE), defined as an event that may have been related to a current medication (e.g., viral infection), but it could not be confirmed. ADEs and PADEs involving either prescription or over-the-counter drugs were included.

22.3.4 Data Collection

Data collection involved a two-step review of ER charts using the Meditech system. Meditech is a hospital information system where all electronic patient information, including ER summaries, are scanned and uploaded to the patient's profile. In the first step, the ER summaries of each selected chart were reviewed by a team consisting of a physician and a registered nurse using a Trigger Assessment Tool. This tool listed 39 screening criteria (triggers) known to be sensitive to the occurrence of ADEs among the adult population. The reviewers combined any triggers found in the ER chart with the patient's history of medication use, as well as a subjective assessment, to determine if an ADE was the reason for the ER visit. If it was classified as being a probable ADE, the reviewers through a consensus process coded the reason for the ER visits as having either a high, moderate, low, or very low probability of being an ADE.

The second step included a full review of all ER charts identified as having "high" and "moderate" probability ADEs, and a random sample of the "low/very low" probability ADEs. As part of the validation exercise, a full review was also carried out on a sample of those ER visits classified as having "no" probability of being ADE. In Step 2, two ER physicians and two clinical pharmacists independently reviewed each of the patient's charts using a data collection tool, which was a modified version of the tool by Gandhi et al. (2003). The reviewers were blinded to the first step review that identified probable ADEs. The reviewers first obtained demographic and clinical information, including presenting complaints, past medical history, drug history, history of allergy, medication dose, frequency, and reaction for the event, as well as the patient's most recent laboratory records.

The reviewers used this information to assess whether the ER visit was a result of an ADE, PADE or ME. Each reviewer also classified the event according to its severity and preventability. Preventability was based on additional information that would have been available had a Pharmacy Network been available. Using an adapted version of previously published criteria (Bates et al., 1995; Gandhi et al., 2003; Gurwitz et al., 2003), severity was classified as being "fatal", "life threatening", "serious" or "significant"; and preventability was classified as "error intercepted", "definitely preventable", "probably preventable", "probably not preventable", or "definitely not preventable". Disagreements about classification of ADEs, and their severity and preventability were resolved during con-

sensus meetings. In this analysis we used two data sets: (a) a limited amount of data collected on all patients from the first review, and (b) detailed information on the subsample of patients that were collected through the chart review.

22.3.5 Statistical Analysis

We generated descriptive statistics including means, standard deviations, and ranges. The primary outcome variables – ADEs and pADEs – were combined into a single variable of ADEs/pADEs in order to reduce the random error associated with the small number of events identified. The unit of analysis was the ER visit. Prevalence of ADEs/pADEs was calculated per 100 ER visits and presented with *p*-values using the binomial proportion test. Each study subject was assigned a sample weight based on the inverse probability of selection. The overall prevalence of ADE/pADEs was estimated using sampling weights to adjust for stratification in the sampling design.

The estimates by age group and sex were kept non-weighted since each patient in the sample frame had an equal chance of being selected within the corresponding age/sex stratum. Events that were assessed as error intercepted, definitely, or probably preventable were merged into one category "preventable", and those assessed to be definitely or probably not preventable were merged into one category "not preventable". The rate of severity and preventability of ADE were derived by dividing the number of events in the respective categories by the total number of ADEs. Mantel-Haenszel chi-square analysis was performed to determine whether there was an association between severity and preventability of ADEs. The number of ADEs was extrapolated to the study population by multiplying the overall prevalence rate by the number of ER visits in the sample frame. Number of preventable ADEs and hospitalization due to ADEs were extrapolated to the study population in a similar manner. All data were entered and stored electronically using Microsoft Access and were analyzed using SPSS 15.0 software package (Statistical Package for Social Sciences, Chicago, IL).

22.3.6 Results

During the study period 82,516 adult ER visits to the HSC and SCMH were identified. Of these, 2,749 visits were excluded because they were by non-residents of N.L. and 12,076 visits were excluded since they did not meet the inclusion criteria, leaving 67,691 ER visits (41,135 unique patients). The mean age (±SD) of this cohort was 46.9 (±19.6) years, with 54.4% (36,814 out of 67,691) of the visits by females. Of the 1,458 ER visits sampled from the 67,691 visits, 44.8% (653) were identified as having a high (29), moderate (135), low (218), or very low (271) probability of being the result of an ADE.

Gastrointestinal symptoms (e.g., nausea, vomiting, and diarrhea) and skin rashes were found to be the most common manifestations of patients identified as high or moderate probability of being an ADE. Patients identified as having a "high" (n = 29) or "moderate" (n = 135) probability of having ADEs, along with a random sample of 170 ER visits classified as having a "low" or "very low" prob-

ability of having ADEs, were independently reviewed by two ER physicians and two clinical pharmacists. The mean (±SD) number of co-morbidities and current medications for this group were 3.5 (± 1.9) and 5.6 (± 3.6), respectively. Fifty-five of the 334 patients were identified by the team to either have an ADE (n = 29) or a PADE (n = 26). After weighting for stratification in the sampling design, the overall prevalence of ADEs/PADEs was 2.8% (95% CI, 2.0-3.7). The mean (±SD) age for patients with ADE/PADE was 69.9 (±14.2); (71.6 ±9.9 for males versus 68.7 ±16.5 for females). No statistically significant difference was found between genders (P = 0.13). For both males and females, the prevalence of ADEs/PADEs increased with age, peaking at 9.1% for females aged 65 years and older. For all age groups, the prevalence of ADEs/PADEs was slightly higher among females than males. In this study, 23 of the 55 patients with ADEs/PADEs (41.8%) required hospitalization.

The mean age for patients with ADEs/PADEs was higher than those having no drug-related visits (69.9 versus 63.8 years, $p < 0.01$). A higher number of co-morbidities and medications were significantly associated with drug-related visits ($p < 0.05$ and $p < 0.01$, respectively). Of the 55 confirmed ADE/PADE patients, one (2%) case was fatal, two (4%) were life-threatening, 25 (46%) serious, and 27 (49%) identified as significant. Approximately 29% of the 55 ADEs/PADEs identified were considered to be preventable had additional information been available through a Pharmacy Network. Of the serious, life-threatening, and fatal events, 35.7% were identified as potentially preventable, compared with 22.2% of the significant events; however, the difference was not statistically significant. Of the 23 hospitalizations due to ADE/PADEs, eight (35%) were considered preventable.

Based on these 55 ADE/PADE patients, we estimate that approximately 1,900 adult patients (95% CI: 1,354-2,505) were treated in the St. John's region for ADEs/PADEs in the two ERs during the study period (January to December 2005), of which an estimated 550 were preventable. Further, of the 1,900 it is estimated that 800 were subsequently hospitalized. This estimate is based on all ER visits (n = 67,691), excluding those not attributed to ADEs (e.g., alcohol-related, suicide attempt, car accidents, cut, burn, wound dressing, etc.). Hematologic complications (e.g., bleeding) were the most common complications associated with ADEs/PADEs (43.6%), followed by gastrointestinal (32.7%), neurological (14.5%), skin (12.7%), cardiovascular (12.7%), metabolic (9.1%), respiratory (7.3%), and renal (5.5%) complications. The medications most frequently associated with ADEs/PADEs, either on their own or in combination with other agents, were such anti-platelets as aspirin (24%), warfarin (18%), antibiotics (15%), anti-hypertensive agents (13%), and chemotherapy agents (11%). Warfarin, divalproex, and chemotherapy agents, medications with a narrow therapeutic index (NTI) and a high risk for toxicity, were found to be the cause of nearly one-third (31.7%) of ER-treated ADEs/PADEs in patients aged 65 years or older. Note that, as part of the validation process, a sample of 192 charts from 805 ER visits classified as "no" probability for ADE visits were reviewed for the validation of the trigger tool exercise. None of these 192 visits were found to be ADE-related.

22.4 Issues, Guidance and Implications

There is considerable research available on ADEs that occur in hospitals, but considerably less so on those that occur in the community. This study is one of the few studies in Canada to investigate ADEs among adult patients presenting to ERs. Our study found that adverse drug events accounted for 2.8% of ER visits, of which about a third were considered preventable if a Pharmacy Network were available. Patients with ADEs/PADEs were found to be older, prescribed more medications, and had a higher number of co-morbidities. Although there is debate in the literature as to whether age itself is a risk factor for an ADE-related visits or hospitalization, the mechanism relating age to risk for ADEs may include the administration of multiple drugs in treating multiple co-morbidities which is more common among the elderly population. In addition, while an aging population tends to take a higher average number of medications, they are also less likely to tolerate certain medications for various reasons, as outlined in the Beers Criteria (Donaldson et al., 2009). In this current study, medications such as warfarin, divalproex, and chemotherapy agents with NTI and high risk for toxicity caused about one-third of ER-treated ADEs in patients aged 65 years or older.

Comparisons with other studies are challenging since there are many variations in case definitions (e.g., ADE, PADE, ME, etc.), study designs, and patient populations. It is argued that the benefit of a Pharmacy Network is its ability to provide a complete patient drug profile on which the pharmacies' drug utilization review software can run, and that that this complete drug profile provides accurate and complete medication information across the continuum of patient care (i.e., Medication Reconciliation). This argument carries significant weight in cases when the patient uses multiple pharmacies when obtaining prescription medications. Conversely, others would argue that where patients only use one pharmacy for all their prescription medications, either out of preference (e.g., knowing the pharmacy staff) or necessity (i.e., the only pharmacy in the community), the benefits of a Pharmacy Network to the patient are minimal.

Another expected benefit of a Pharmacy Network is the reduction in double doctoring, as prescription-dispensing records would be available to all pharmacies on the network in real time. The other issue sometimes raised is that there is a usually a cost to the pharmacy for being part of the network (e.g., hardware and software upgrades, Internet access, lost productivity, etc.) and that the pharmacy is a private company that for the most part generates revenue through dispensing medications, not providing additional patient care. While there can be several valid arguments, both for and against a Pharmacy Network, ultimately if it provides increased patient safety and improves quality of care, both government and the private sector need to work towards the deployment of such a network across their population.

This study faced several limitations. Firstly, using a retrospective chart review design may underestimate the true frequency of emergency visits as being caused by an ADE. Ideally, a prospective design with a large sample including patient interviews and obtaining key information would have increased the ac-

curacy of estimates of drug-related visits and their preventability. Secondly, compared to patients aged 65 years or more, we found fewer ADEs in younger age groups, which makes our estimates of ADE prevalence more prone to sampling error in these age groups. Nevertheless, the prevalence of ADE among elderly patients was 8.4%, which is very close to our pre-study assumption of 10% considered in the study design.

Thirdly, in preparing the sampling frame, we excluded 12,076 visits from the study population using pre-defined exclusion criteria. However, based on a cursory review of these excluded visits, we concluded that the criteria used in excluding non-drug related visits might not have been as precise as we had hoped. The main reason for this lack of precision was that the exclusion criteria were applied to the patients' self-reported complaint, and not the diagnosis provided by the attending health professional after the encounter. As such, exclusion of any drug-related visits may have resulted in leaving out a low-risk subset of ER visits instead of a no-risk subset, and thereby resulted in overestimating the prevalence of ADEs in the study sample. Fourthly, we did not extrapolate our data to the entire province, since the HSC and SCMH are located in the capital city of St. John's and cannot be considered representative of all hospitals in the province.

22.4.1 Summary of Evaluation Issues

The evaluation of a Pharmacy Network presented issues that exist with most evaluations, in that there is a lack of standards in undertaking evaluations overall, which limits the amount of comparability one study has. The methodological approach to evaluations is not new, with most employing age-old research methods (e.g., surveys, interviews, chart reviews, administrative data, etc.) to determine whether whatever is being evaluated has met its objectives. The challenge is getting a consistent approach so that peer-to-peer comparisons can be made and best practices identified. In the absence of such comparisons, we are limited to comparing results in the same environment pre- and post-implementation, with no idea if the pre-intervention indicators are any better (or worse) than our peers. In the case of this current study, the team is waiting until the Pharmacy Network has been fully deployed for 12 months before doing the post-Pharmacy Network intervention. This is expected to occur in early 2018.

22.4.2 Guidance for Future Directions

Evaluating the benefits of a Pharmacy Network is not only resource intensive and costly, but is delivered within a government's policy framework and as such is not under the control of the evaluation team. When evaluating a government intervention, whether it is a policy, a program, or a new technology, always consider that many issues that will arise will be out of your control and you must mitigate them as best you can in the design of your evaluation.

22.4.3 Policy and Practice Implications

In implementing a Pharmacy Network it is in the interest of government to provide its population with a sustainable, high quality, and safe service in relation to the usage of prescription medications. Through that lens it seems logical that a Pharmacy Network would deliver on these three fronts, ignoring the costs to actually implement the network. However, in the practice environment it is not so linear, as some pharmacies may not perceive any benefits if they believe their client population is non-nomadic. If a Pharmacy Network does not include all pharmacies within the population, health professionals may not be provided with their patient's complete drug profile, reducing double doctoring is compromised, and the data will be incomplete in the development of new policies and programs.

22.5 Summary

Emergency room visits as a result of ADEs are not uncommon. A focus on further education along with the tools need to be in place so that physicians and pharmacists can collaborate more closely to improve prescribing practices and monitoring, particularly among high-risk patients, and thereby contribute to reducing the subset of ADEs that is potentially preventable. The authors believe that if a Pharmacy Network were deployed it would allow authorized healthcare providers to access and share information, which would contribute to reducing the frequency of adverse events related to drugs in the community.

References

Alexopoulou A., Dourakis, S. P., Mantzoukis, D., Pitsariotis, T., Kandyli, A., Deutsch, M., & Archimandritis, A. J. (2008). Adverse drug reactions as a cause of hospital admissions: A 6-month experience in a single center in Greece. *European Journal of Internal Medicine, 19*(7), 505–510.

Baker, G. R., Norton, P. G., Flintoft, V., Blais, R., Cox, J., Etchells, E., … Tamblyn, R. (2004). The Canadian adverse events study: the incidence of adverse events among hospital patients in Canada. *Canadian Medical Association Journal*, *170*(11), 1678–1686.

Bates, D. W. (1998). Drugs and adverse drug reactions: How worried should we be? *Journal of the American Medical Association, 279*(15), 1216–1217.

Bates, D. W., Cullen, D. J., Laird, N., Petersen, L. A., Small, S. D., Servi, D., … Edmondson, A. (1995). Incidence of adverse drug events and potential adverse drug events. *Journal of the American Medical Association, 274*(1), 29–34.

Bates, D. W., Spell, N., Cullen, D. J., Burdick, E., Laird, N., Petersen, L. A., … Leape, L. L. (1997). The costs of adverse drug events in hospitalized patients. Adverse drug events prevention study group. *Journal of the American Medical Association, 277*(4), 307–311.

Budnitz, D. S., Pollock, D. A., Mendelsohn, A. B., Weidenbach, K. N., McDonald, A. K., & Annest, J. L. (2005). Emergency department visits for outpatient adverse drug events: Demonstration for a national surveillance system. *Annals of Emergency Medicine*, *45*(2), 197–206.

CBC News. (2015, September 29). More Canadians are 65 and over than under age 15, StatsCan says. *Canadian Broadcasting Corporation News/Business*. Retrieved from http://www.cbc.ca/news/business/statistics-canada-seniors-1.3248295

Canadian Institute for Health Information. (2010). *Drug use among seniors on public drug programs in Canada, 2002-2008*. Ottawa, ON: Author.

Chrischilles, E. A., VanGilder, R., Wright, K., Kelly, M., & Wallace, R. B. (2009). Inappropriate medication use as a risk factor for self-reported adverse drug effects in older adults. *Journal of the American Geriatrics Society*, *57*(6), 1000–1006.

Donaldson, M. G., Sobolev, B., Cook, W. L., Janssen, P. A., & Khan, K. M. (2009). Analysis of recurrent events: a systematic review of randomised controlled trials of interventions to prevent falls. *Age and Ageing, 38*(2), 151–155. doi: 10.1093/ageing/afn279.

Ernst, F. R., & Grizzle, A. J. (2001). Drug-related morbidity and mortality: updating the cost-of-illness model. *Journal of the American Pharmaceutical Association (Washington), 41*(2), 192–199.

Evans, R. S., Lloyd, J. F., Stoddard, G. J., Nebeker, J. R., & Samore, M. H. (2005). Risk factors for adverse drug events: a 10-year analysis. *Annals of Pharmacotherapy*, *39*(7), 1161–1168.

Fialová, D., Topinková, E., Gambassi, G., Finne-Soveri, H., Jónsson, P. V., Carpenter, I., … Bernabei, R. (2005). Potentially inappropriate medication use among elderly home care patients in Europe. *Journal of the American Medical Association, 293*(11), 1348–1358.

Field, T. S., Gurwitz, J. H., Avorn, J., McCormick, D., Jain, S., Eckler, M., … Bates, D. W. (2001). Risk factors for adverse drug events among nursing home residents. *Archives of Internal Medicine, 161*(13), 1629–1634.

French, D. G. (1996). Avoiding adverse drug reactions in the elderly patient: issues and strategies. *Nurse Practitioner, 21*(9), 90–105.

Gandhi, T. K., Weingart, S. N., Borus, J., Seger, A. C., Peterson, J., Burdick, E., … Bates, D. W. (2003). Adverse drug events in ambulatory care. *New England Journal of Medicine, 348*(16), 1556–1564.

Glynn, R. J., & Buring, J. E. (1996). Ways of measuring rates of recurrent events. *British Medical Journal, 312*(7027), 364–367.

Gurwitz, J. H., Field, T. S., Harrold, L. R., Rothschild, J., Debellis, K., Seger, A. C., …Bates, D. W. (2003). Incidence and preventability of adverse drug events among older persons in the ambulatory setting. *Journal of the American Medical Association*, *289*(9), 1107–1116.

Health Canada. (2010). Adverse reaction and incident reporting–2009. *Canadian Adverse Reaction Newsletter, 20*(2), 2–4.

Institute of Medicine. (1999). *To err is human: Building a safer health system*. Washington, DC: National Academy Press.

Juntti-Patinen, L., & Neuvonen, P. J. (2002). Drug-related deaths in a university central hospital. *European Journal of Clinical Pharmacology, 58*(7), 479–482.

Johnson, J. A., & Bootman, J. L. (1995). Drug-related morbidity and mortality. A cost-of-illness model. *Archives of Internal Medicine, 155*(18), 1949–1956.

Kaufman, D. W., Kelly, J. P., Rosenberg, L., Anderson, T. E., & Mitchell, A. A. (2002). Recent patterns of medication use in the ambulatory adult population of the United States: the Slone survey. *Journal of the American Medical Association*, *287*(3), 337–344.

Lazarou, J., Pomeranz, B. H., & Corey, P. N. (1998). Incidence of adverse drug reactions in hospitalized patients: a meta-analysis of prospective studies. *Journal of the American Medical Association*, *279*(15), 1200–1205.

McCaig, L. F., & Burt, C. W. (2003). *National hospital ambulatory medical care survey: 2001 emergency department summary* (Advance Data from Vital and Health Statistics, No. 335). Hyattsville, MD: National Center for Health Statistics, Department of Health and Human Services.

McDonnell, P. J., & Jacobs, M. R. (2002). Hospital admissions resulting from preventable adverse drug reactions. *Annals of Pharmacotherapy*, *36*(9), 1331–1336. doi: 10.1345/aph.1A333

Morimoto, T., Gandhi, T. K., Seger, A. C., Hsieh, T. C., & Bates, D. W. (2004). Adverse drug events and medication errors: detection and classification methods. *Quality and Safety in Health Care, 13*(4), 306–314.

Nebeker, J. R., Barach, P., & Samore, M. H. (2004). Clarifying adverse drug events: a clinician's guide to terminology, documentation, and reporting. *Annals of Internal Medicine, 140*(10), 795–801.

Nebeker, J. R., Hoffman, J. M., Weir, C. R., Bennett, C. L., & Hurdle, J. F. (2005). High rate of adverse drug events in a highly computerized hospital. *Annals of Internal Medicine, 165*(10), 1111–1116. doi: 10.1197/jamia.M155

Onder, G., Landi, F., Vedova, C., Atkinson, H., Pedone, C., Cesari, M., Bernabel, R., & Gambassi, G. (2002). Moderate alcohol consumption and adverse drug reactions among older adults. *Pharmacoepidemiology and Drug Safety, 11*(5), 385–392.

Onder, G., Penninx, B. W., Landi, F., Atkinson, H., Cesari, M., Bernabel, R., & Pahor, M. (2003). Depression and adverse drug reactions among hospitalized older adults. *Archives of Internal Medicine*, *163*(3), 301–305.

Pirmohamed, M., James, S., Meakin, S., Green, C., Scott, A. K., Walley, T. J., … Breckenridge, A. M. (2004). Adverse drug reactions as cause of admission to hospital: prospective analysis of 18,820 patients. *British Medical Journal, 329*(7456), 15–19.

Roughead, E. E., Gilbert, A. L., Primrose, J. G., & Sansom, L. N. (1998). Drug-related hospital admissions: a review of Australian studies published 1988-1996. *Medical Journal of Australia, 168*(8), 405–408.

Samoy, L. J., Zed, P. J., Wilbur, K., Balen, R. M., Abu-Laban, R. B., & Roberts, M. (2006). Drug-related hospitalizations in a tertiary care internal medicine service of a Canadian hospital: a prospective study. *Pharmacotherapy*, *26*(11), 1578–1586.

Statistics Canada. (2007, Winter). Census snapshot of Canada – Population (age and sex). Catalogue 11.008. *Canadian Social Trends, 84*, 37–38. Ottawa: Author.

Tafreshi, M. J., Melby, M. J., Kaback, K. R., & Nord, T. C. (1999). Medication-related visits to the emergency department: a prospective study. *Annals of Pharmacotherpy, 33*(12), 1252–1257.

Zed, P. J., Abu-Laban, R. B., Balen, R. M., Loewen, P. S., Hohl, C. M., Brubacher, J. R., … Purssell, R. A. (2008). Incidence, severity and preventability of medication-related visits to the emergency department: a prospective study. *Canadian Medical Association Journal*, *178*(12), 1563–1569.

Zhang, M., Holman, C. D. J., Price, S. D., Sanfilippo, F. M., Preen, D. B., & Bulsara, M. K. (2009). Comorbidity and repeat admission to hospital for adverse drug reactions in older adults: Retrospective cohort study. *British Medical Journal, 338*(a2752), 1–9. doi: 10.1136/bmj.a2752

Chapter 23

Evaluation of Electronic Medical Records in Primary Care

A Case Study of Improving Primary Care through Health Information Technology

Lynne S. Nemeth, Francis Lau

23.1 Introduction

Translating research into practice continues to be a challenge within primary care, evidenced by the inconsistent performance in delivery of recommended healthcare (McGlynn et al., 2003). Organizational support for improvement and implementation of guideline-based care in small primary care practices, where the majority of healthcare is delivered, is often fragmented and underdeveloped (Fisher, Berwick, & Davis, 2009).

Workflows in the primary care office are sometimes complicated and inefficient, and replacing paper records with an EMR system does not fix these inefficiencies (Miller & Sim, 2004). To improve quality when implementing the EMR, workflow redesign is important (Fiscella & Geiger, 2006). Smaller primary care practices that operate outside of large healthcare systems often lack systematic resources that assist them to set priorities for quality improvement (QI), and develop the staff that provide and support clinical care. Time and resources are needed to address the steep learning curve and the knowledge development needs of the non-clinician staff. A flexible change management strategy (Lorenzi, Kouroubali, Detmer, & Bloomrosen, 2009) and strategic planning is needed when adopting electronic medical records that can be used beyond the ways of a paper medical record system (Baron, 2007).

HIT adoption may be the catalyst to stimulate a process of change in the provision of primary care service delivery that improves the practice as a system.

When all members of the primary care team have timely access to patient information, the overall coordination of care can be improved, and team members can take on new roles that enhance the quality of healthcare. Yet using EMRs, clinical decision support systems, order entry, appointment schedules, and test results reporting systems requires the adoption and the creation of best practices in implementation, use and maintenance of the systems. Transformation of primary care is needed to redesign the system for improved care coordination, quality and safety, which benefit from better use of HIT (Meyers & Clancy, 2009). A coherent model to assist practices with implementation of evidence-based guidelines using HIT, grounded in the real-world experiences of small primary care offices, can engage newer practices on the path to improve the quality and effectiveness of the healthcare delivered.

23.2 Case Study: Improving Primary Care Through Health Information Technology

23.2.1 Aims

A series of seven studies that focused on Translation of Research into Practice (TRIP) within the Practice Partner Research Network (PPRNet), a primary care practice-based research network in the United States, were selected for secondary analysis to synthesize a decade of learning regarding how to use health information technology (HIT) to improve quality in primary care practices. A comparative case analysis of the findings of the seven studies created new insights regarding improving quality using HIT. The specific aims of this project were:

1 Complete a mixed methods secondary analysis to synthesize findings on using information technology (IT) to improve quality in primary care across seven nationally funded PPRNet initiatives.

2 Examine current perspectives of PPRNet-TRIP participants on team development and on methods for developing and sustaining QI efforts.

3 Integrate findings from PPRNet's previous studies with the current perspectives of practice representatives to refine the overarching theory-based "PPRNet-TRIP QI Model".

23.2.2 Background, Context, Settings, and Participants

PPRNet has conducted a consecutive series of research studies focused on TRIP funded by several United States agencies (Agency for Healthcare Research and Quality [AHRQ], National Cancer Institute [NCI], and National Institute for Alcoholism and Alcohol Abuse [NIAAA]) since 2001. This national practice-

based research network was established in 1995, at the Medical University of South Carolina. Up to 225 practices from 43 states in the U.S. participated in PPRNet activities through the quarterly extraction of electronic health record data from their practices, for benchmarking and quality improvement, and participation in PPRNet research trials and demonstration projects awarded during this time. Network participants shared best practices in improving quality on selected areas of interest at annual network meetings convened to form a collaborative learning community hosted by PPRNet investigators. This particular study was implemented to reach across a body of research that had focused on specific clinical areas for improvement, to generate overarching lessons learned from a decade of specific research that translated research into practice using electronic health records (EHRs).

23.2.3 Methods (Study Design, Data Sources/Collection)

Aim 1:
The mixed methods data from seven PPRNet studies were merged into an NVivo 9.0 database for qualitative secondary analysis. The studies focused on the following indicators and were funded by the following agencies.

- TRIP-II (cardiovascular disease and stroke secondary prevention) *AHRQ*
- A-TRIP (36 primary care indicators) *AHRQ*
- AA-TRIP (alcohol screening, brief intervention) *NIAAA*
- C-TRIP (colorectal cancer [CRC] screening) *NCI*
- MS-TRIP (medication safety) *AHRQ*
- SO-TRIP (screening, immunizations and diabetes care management) *AHRQ*
- AM-TRIP (alcohol screening, brief intervention, medication) *NIAAA*

Data were incorporated from the variety of sources from the participation of 134 practices (e-mail, meeting notes, site visit evaluations, focus groups, interviews, observations, memos) for analyses within the NVivo 9.0 database. Additionally, the performance data on PPRNet measures were reviewed to identify practices that were effective in implementing changes to improve performance in their practices on selected measures. In the review of these various data, concepts related to how practices revised clinical processes, procedures

and roles were clarified and compared across studies. Practice strategies for improvement within practices were examined after intense immersion with the data, and a cross-case comparison method enabled discovery of common features of each of the cases. Each of the studies listed above was considered a case. An inductive and deductive process was used iteratively in coding the data. The aim of these analyses were to draw out new ideas, to expand on concepts previously noted in these studies, and also to fit data into categories representing newer strategies that evolved over the decade. Current literature representing the advances over the decade in HIT, quality and patient-centred care were used to search more deductively for evidence of characteristics of these trends. An emphasis on data reduction was needed to minimize redundancy/overlap in concepts and to improve the clarity of a model for improvement that might be used to develop practices that are newer in their adoption of HIT.

Aim 2:
The 2011 and 2012 PPRNet annual meetings provided opportunities to review the current perspectives of PPRNet-TRIP participants. These diverse, national audiences of PPRNet practice members participated in the meetings held in Charleston, South Carolina for networking and dissemination of best practices related to Medication Safety, Standing Orders, Alcohol Screening and Brief Intervention, and Judicious Use of Antibiotics for Acute Respiratory Infections. Participants represented rural, urban, community-based family and internal medicine practices and included clinicians, clinical staff, practice managers, HIT support staff, and other office staff, primarily from small- to medium-sized practices, but including a number of larger practices as well. Field notes were taken regarding the Medication Safety component of the 2011 meeting that reflected how practices that participated in the MS-TRIP 2 project made improvements in their practice, why working on medication safety mattered to them, case examples of best practice strategies, and how these improvements related to efforts towards Patient Centred Medical Home (PCMH), meaningful use, and other aspects of performance review. Practices shared their best practice plans, and discussed timelines for implementing these plans.

A theme of the 2011 annual meeting focused on using PPRNet reports and quality improvement approaches to achieve Patient Centred Medical Home (PCMH) and other quality recognitions. One of the specific components of the 2011 meeting included a presentation of "Lessons Learned from 10 years of Translating Research into Practice" (Nemeth, 2011), and a panel of practice staff and providers from four practices that had exemplified numerous strategies that were learned from *Aim 1*. The practice panel provided an opportunity to seek the perspectives of other practices on how team development and sustaining quality improvement occurred in practice. Field notes were collected at this meeting (the 2011 meeting included 113 participants, with 57 practices represented) to document the discussion. Topics included: practice progress towards improving quality through participation in PPRNet; what has evolved and im-

proved; how this was accomplished; and what is most important to develop a team practice, to adopt and use HIT tools, to transform practice culture and quality, and to activate patients. There was a deep review of concepts, discussion of strategies and many questions and dialogue from the meeting participants, including discussion of potentially missing components from the model.

An interview guide was pretested with four practices, and these four practices presented their views on practice development and sustainability for QI at a panel presentation. Telephone interviews followed up the annual 2011 meeting to gain perspectives of other providers and staff that had participated in PPRNet research. Interviews were conducted between 2011 and 2012, in the context of current research underway within each practice, or practice initiatives to improve and capture additional practice revenue from payer initiatives, such as Patient Centred Medical Home pilots or Meaningful Use. The practice activities underway during the years 2010 through 2012 incorporated new interests in incentives with healthcare reform legislation passed.

The 2012 annual meeting included 98 practice participants, from 46 practices. In a session related to promoting the judicious prescribing of antibiotics for acute respiratory infections, we gathered practice perspectives in field notes related to the use of a template for clinical decision support, how to embed patient education into a structured visit guided by a template, and how to respond to patients requesting antibiotics when they were not indicated. Regarding the AM-TRIP project, we collected practice comments regarding the use of alcohol screening and brief interventions, and medication management for high-risk drinkers. The discussion reflected challenges with patients, reluctance from providers and nursing staff and how these were overcome in practices that participated in this study. Practice participants who did not participate in this study had the opportunity to learn from these practices, and raise awareness of the progress of other participating practices in improving performance on alcohol screening, intervention and treatment. Field notes taken during these sessions documented additional perspectives to the qualitative data that underlies the refined model.

Aim 3:

This aim involved a creative synthesis in mapping the key concepts as variables that impact the process of improving primary care. Once the four key concepts were identified, the inputs and outputs related to these activities were mapped as a visual logic model. Yet, the visual representation of the relationships between these concepts evokes an understanding that makes practical sense to many practicing clinicians and their staff who provide primary care. After developing the visual figure the concepts and model were reviewed and, after an iterative process of revision and presentation to numerous audiences, the new PPRNet model was finalized. A logic model was added to more clearly specify how the model can be used as an implementation and evaluation framework similar to other implementation science efforts.

23.2.4 Results (Principal Findings, Outcomes, Discussion, Conclusions, Significance, Implications)

Aim 1. Secondary analysis of seven studies. The original PPRNet-TRIP QI model was developed through grounded theory development in the TRIP-II and A-TRIP studies which were formative to the subsequent PPRNet body of research. It became clear after lengthy immersion in the data, reflecting on the evolution of practice activities over the decade, that greater sophistication about how to improve on quality measures had occurred, and that many practices were highly motivated to achieve a competitive position. Four main concepts central to the new framework were identified: (a) developing a team care practice; (b) adapting and using HIT tools; (c) transforming the practice culture and quality; and (d) activating patients. The four concepts emphasize the complex interactions and roles within primary care practice, and interventions related to improvement on performance measures. Figure 23.1 presents the framework, and Figure 23.1 elaborates how the concepts in the early studies led to more sophisticated and complex practice transformation.

Aim 2: Examine current perspectives of PPRNet-TRIP practice participants on team development and on methods for sustaining QI efforts. Twenty interviews were conducted with primary care providers of practices in PPRNet after the development of the revised model. The findings of these interviews contributed to furthering an understanding of how practices developed their teams, and what enabled them to sustain their efforts to improve. These interviews elaborated provider perspectives about how they have developed during a more recent trend towards rewards for quality and performance in ambulatory care, a desire for designation as patient-centred medical homes, and participation in early pilots from commercial payers, Medicare and Medicaid demonstration projects, and meaningful use.

The key perspectives included support for developing enhanced roles for staff in the practice to collect more data from patients, acting on decision support, reminders, and alerts provided within the EHR, and implementing routine actions that save the provider time during clinical encounters. The need for technical support to ensure that the EHR was set up correctly to provide the needed health information to be alerted was clearly articulated, and often the role of technical support was provided by a lead physician who was more technically savvy than others or more inclined to take on this responsibility. In practices that lacked this internal leadership, and in larger practices, IT support staff were needed and worked with a lead provider. Care coordination and outreach to follow up on patients not at goals for values of quality measures, or for those that needed chronic care management, was clearly becoming a more important activity in practices that wanted to act on the performance data that was generated within PPRNet reports. The activities related to increasing patient-centredness and patient activation were newer activities in many of the practices, and the EHR resources proved to be a very important component of reaching out to patients using Web portals, letters, and after-visit summaries and re-

minders to patients to follow up on issues that were important to their care. Most of these additional activities were undertaken to reap financial rewards for the quality of care that the practice was aiming for.

The interviews established validity for the revisions to the PPRNet-Translating Research into Practice (TRIP) QI model that had been used within practices to improve quality of care using HIT.

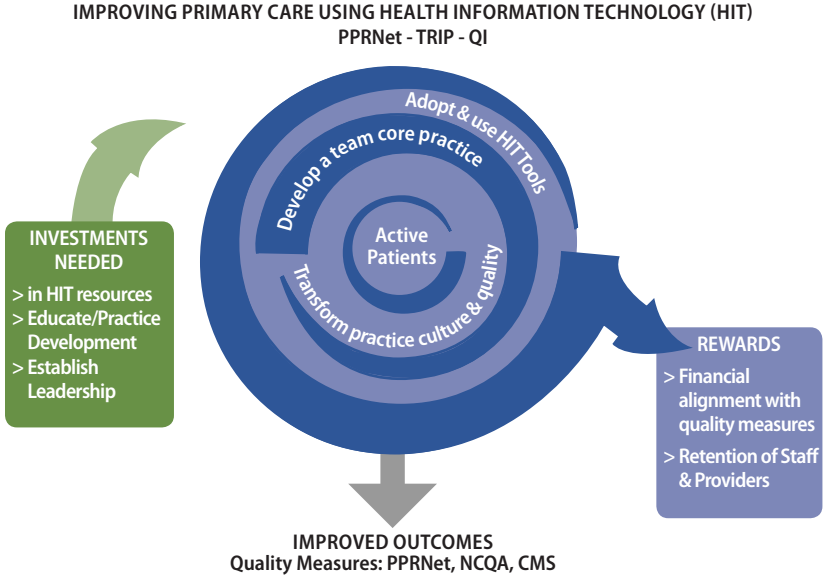


Figure 23.1. PPRNet-TRIP QI: A refined framework guiding primary care improvement.

Note. From "Lessons learned from 10 years of translating research into practice," by L. Nemeth, 2011, a presentation to the 16th annual meeting of the Practice Partner Research Network (PPRNet), Medical University of South Carolina, Charleston, SC. [Acknowledgement: AHRQ R03HS018830].

Aim 3: Integrate findings from PPRNet's previous studies with the current perspectives of practice representatives to refine the overarching theory-based "PPRNet-TRIP QI Model". The four concepts in the new model — "Improving Primary Care through Health Information Technology" (IPC-HIT) — provide clear areas of focus for developing primary care practices towards high performance on quality measures using HIT. Figure 23.1 presents the concepts and relationships of the framework. The inputs to the process viewed as a logic model include that practices must decide to make investments in HIT resources, which require the financial capacity and time to be allocated for selection and learning to use the EHR. Education of the providers and staff is required, and leaders must be appointed to ensure appropriate use of new systems. In some practices this required hiring HIT coordinators, and in others a technically savvy clinician might take the lead. Outputs of the process shown in the centre of the model in

the figure include: (a) financial rewards to the practice for their accomplishments in improvement, and (b) retention of staff and providers who work together to increase value in the healthcare services provided. Outcomes are demonstrated performance improvements on measures that are important to the practice, such as PPRNet quality measures, and how they stand on these measures compared to the other practices in PPRNet as noted by PPRNet medians and benchmarks (90th percentile).

Primary care practices that have used EHRs, participated in PPRNet practice-based research to improve the translation of research into practice, or have been willing to share their strategies, successes, barriers, and rewards have been able to make improvements towards higher performance. This learning community has provided opportunities for reciprocal knowledge dissemination from researchers to clinicians and vice versa. The lessons of this decade of research together provide a model for other practices newer in the transition and adoption of EHR tools to improve quality using their enhanced teams and a quality culture to activate patients.

To explain these concepts in more detail, Table 23.1 presents "what" (concepts) and explains "how" (strategies) improvements in primary care have been made during participation in PPRNet studies.

The logic model for IPC-HIT is presented in Table 23.2. For practices that are implementing the IPC-HIT model the following strategies and measures should be considered.

Table 23.1
Specific Approaches Found Within a Decade of PPRNet Research

		Improving Primary Care Using HIT: Specific Approaches/Strategies Within PPRNet		
		TRIP-II to ATRIP (2001-2006)	**AA/AM/SO/C-TRIP (2005-2012)**	**MS-TRIP (2007-2012)**
Concepts	Develop a Team Care Practice	• "Involve all staff," new roles/responsibilities • Clinicians agree to decrease practice variation	• Structured screening tools (MAs/nurses) • Complementary team roles better defined, providers closing loop	• Medication reconciliation, outreach as needed
	Adapt and Use HIT Tools	• Staff increased use of EHR	• Specific templates used for decision support • Revised/edited, add macros, applied age, gender, Dx/ Rx templates • Lab interfaces, scanning, eRX, Web-based patient portals added	• Rx/Dx templates applied • Improved medication reconciliation • Increased attention to dosing alerts
	Transform Practice Culture and Quality	• Emphasis on quality, set goals, celebrated successes • Quality committees/ coordinators	• Liaisons coordinate projects/communication, use performance reports at practice and patient level • Staff education; SO's increased, explicit policies, practice culture rewarded by P4P etc.	• Reports used for outreach • Refill protocols • Standing orders for labs • Printed med lists
	Activate Patients	• Handouts, posters, screening/immunization events • Press releases	• Brief intervention, counselling, treatment, referrals • Targeted messages: "Rethinking Drinking"; Screen for Life; birthday letters, HM reminders in letter • Active f/u for completion of tests; outreach	• Patient update forms, bring all meds, labs in advance • Long appointments for med reviews, med list provided at end of visit

Note. From "Making sense of electronic medical record adoption as complex interventions in primary health care," by F. Lau, L. Nemeth, and J. Kim, 2012, a presentation to the *Canadian Association for Health Services and Policy Research (CAHSPR)* conference, Montreal. [Acknowledgement: AHRQ R03HS018830].

Table 23.2
Logic Model Disseminating Effective Strategies to Improve Preventive Services Using HIT

Construct	Practice Strategy	Measures of Implementation /Outcomes
Develop a Team Care Practice	• Design practice roles and processes that support workflow • Provide tools to clarify care process, staff understand new roles • Create environment of mutual trust and open communication • Recruit staff members comfortable working in an empowered practice • Regular team meetings to determine best processes	• Process evaluation (Q): - Roles clear/adopted - Policies/protocols - Communication mechanisms in place - Staff selected that embrace practice goals/retained - Team meets regularly and engaged in decision-making
Adapt and Use Health Information Technology Tools	• Leader oversees adapting and updating electronic health record (EHR) for clinical decision support • Use embedded utilities to ensure age, gender and condition specific templates are applied within EHR • Embed structured templates for staff data collection and follow-up • Use medication prescribing alerts and e-prescribing • Interface labs, scan procedure reports and outside services to ensure accurate records	• Time and cost allocated for HIT support by practice (who, how much time, financial impact) (S) • Extent of use of CDS tools among practice staff and providers (S) • Proportion of patients within practice with e-prescriptions (PR) • Proportion of patients with up-to-date health maintenance (HM) received (PR)
Transform Practice Culture and Quality	• Review performance reports to identify priorities for improvement • Practice-wide discussion and agreement re: quality goals • Training to increase staff self-efficacy to implement changes • Evaluate and support learning and improvement efforts as a team	• Performance on selected quality measures improved (PR) • Staff adopt roles/responsibilities (S) • Providers perceive effectiveness of workflow (S) • Practice receives increased revenues for performance (S)
Activate Patients	• Engage patients through screening conversations and reminders • Use posters, letters, and Web portals • Outreach to ensure completion of recommended services • Remind patients to bring all medications to visits for; reconciliation and review	• Process evaluation to assess (Q): - Posters, letters - Patient Web portals, kiosks - Review and reconciliation processes - Outreach • Up to date HM received (PR)
Legend: Measures can be evaluated by: Qualitative data (Q); Performance Reports (PR); or Survey (S)		

Acknowledgement: AHRQ R03HS018830 (Final progress report to AHRQ, 2013, unpublished)

23.2.5 Issues, Guidance and Implications

Developing a Team Care Practice adds an understanding that providers engage their staff as partners to achieve quality outcomes with patients. Front office staff members who receive and schedule patients for follow-up care need to fully understand the goals for improvement that the practice has set, to increase the follow through by patients. Clinical staff participated to a larger extent in role expansion when they were clear about the goals, what the practice wanted to do as a team, and knew what their role expectations were.

Adapting and Using Health Information Tools involves developing more sophisticated use of the HIT tools available in the practice's EHR. Effective use of EHR features requires practice customization related to patient populations served, and practice patterns. Some degree of HIT expertise is needed to be able to customize these tools to provide efficient and accurate data that drives reminders, alerts and any other decision support that is needed to deliver quality primary care. This may require the allocation of practice-based resources to ensure this component is managed effectively.

Transformation of Practice Culture and Quality is a process that evolves from engagement as a team, and using data from performance to inspire practices to develop new approaches. This occurs while learning, evaluating and reflecting on practice-specific progress in the improvement efforts that have been prioritized, and the research evidence that has been translated into practice.

Lastly, *Activate Patients* is the focus of practice-based efforts to improve. This often was seen as a paradigm shift from an era of provider-dominated healthcare agendas to a focus on developing patient-centredness in an era of stakeholder-engaged teams seeking to improve knowledge regarding healthcare decisions and behaviours, activation of patients as partners in their care, and understanding of values and preferences of patients. In this study we learned that by using HIT tools, practice teams can reach out to patients to provide and validate recorded health information data, present needed services, request patient decisions and ensure medications are reconciled, and monitor chronic conditions as needed.

Noted within this synthesis of seven studies were both barriers and facilitators to improvement in primary care using HIT.

1 Barriers included: lack of practice leadership, vision and goals related to improvement using HIT; lack of provider agreement and consensus on approaches; need for HIT technical support, expertise and resources for using HIT effectively; staff and provider turnover, organizational change or change in practice ownership.

2 Facilitators included: having practice policies and protocols; staff education and follow-up by leaders and clinicians; enhanced communication processes; streamlined tools and templates to improve workflow and efficiency; having a practice-wide approach that re-

inforced consistent staff expectations for adoption of expanded roles; and having providers close the loop on what practice staff initiate.

New questions and hypotheses were generated by this research. Most importantly, the introduction of the four concepts in the IPC-HIT model provide direction to practices that want to improve their workflow and processes to achieve goals of improved healthcare delivery to their activated patients. By introducing the concepts and example practice strategies for improving primary care through HIT, there should be corresponding implementation plans and measurement of outcomes such as noted in the logic model in Table 23.2. Some examples of the hypotheses related to processes and outcomes found in this table include:

- Staff will adopt expanded roles with clear policies and protocols regarding using the HIT in their work with patients.
- Providers will close the loop with patient care when staff members initiate patient services that are warranted by practice protocol.
- HIT will be supported by a designated leader within the practice, who will educate staff and providers regarding changes.
- Performance on clinical quality measures show improvement after developing practice teams with this model.
- Financial revenues are increased related to performance on clinical quality measures.
- Providers and staff are retained in practices that provide attention to the four concepts in the model.

A primary limitation to this research should be noted. The principal investigator was the qualitative analyst of the original research and this synthesis. Limited resources to review the wealth of qualitative data obtained in the primary studies precluded analytic support. However, with the assistance of the primary researchers, and review by the member practices in PPRNet, it was clear that the model was supported as valid. Overcoming this limitation, it should be noted that the strength of the research was that it was conducted in a national network and not limited to a specific geographic region. Participants in PPRNet were clear about how they develop their staff toward high performance, and have a track record evidenced in their performance data that demonstrated the effective approaches resulted in clear improvements.

23.3 Summary

Over the past decade, PPRNet established a theoretically-informed framework for translating research into practice (TRIP) in small- to medium-sized primary care practices that use the Practice Partner® electronic medical record (EMR). The PPRNet-TRIP Quality Improvement (QI) Model included three components: an intervention model, an improvement model, and a practice development model that assists practices with implementation of strategies to improve on selected performance measures. During the course of the present research, we have streamlined the most important components to four main concepts that can provide an organizing framework for improvement.

This research included a robust evaluation of the mixed methods data and lessons learned from a decade of PPRNet-TRIP. The experience of PPRNet research participants and researchers enhanced understanding of the PPRNet-TRIP components and how practices improve primary care quality with their health information technology and team-based approaches to care. The cross-case analyses conducted through this research generated important themes, provided new insights, and generated new hypotheses about factors that improve the quality of care through the use of EMRs. The new framework will provide practical guidance for practices that are undertaking these efforts to achieve meaningful use, patient centred medical home recognition and paths for improved financial resources pertaining to quality improvement in primary care practice.

References

Baron, R. J. (2007). Quality improvement with an electronic health record: Achievable, but not automatic. *Annals of Internal Medicine, 147*(8), 549–552.

Fiscella, K., & Geiger, H. J. (2006). Health information technology and quality improvement for community health centers. *Health Affairs, 25*(2), 405–412. doi: 10.1377/hlthaff.25.2.405

Fisher, E. S., Berwick, D. M., & Davis, K. (2009). Achieving health care reform — How physicians can help. *New England Journal of Medicine, 360*(24), 2495–2497. doi: 10.1056/NEJMp0903923

Lau, F., Nemeth, L. S., & Kim, J. (2012, May 30). *Making sense of electronic medical record adoption as complex interventions in primary health care.* Presentation to the Canadian Association for Health Services and Policy Research (CAHSPR) Conference, Montreal.

Lorenzi, N., Kouroubali, A., Detmer, D., & Bloomrosen, M. (2009). How to successfully select and implement electronic health records (EHR) in small ambulatory practice settings. *BMC Medical Informatics and Decision Making, 9*(1), 15.

McGlynn, E. A., Asch, S. M., Adams, J., Keesey, J., Hicks, J., DeCristofaro, A., & Kerr, E. A. (2003). The quality of health care delivered to adults in the United States. *New England Journal of Medicine, 348*(26), 2635–2645.

Meyers, D. S., & Clancy, C. M. (2009). Primary care: Too important to fail. *Annals of Internal Medicine, 150*(4), 272–273.

Miller, R. H., & Sim, I. (2004). Physicians' use of electronic medical records: barriers and solutions. *Health Affairs, 23*(2), 116–126.

Nemeth, L. S. (2011, August 21–23). *Lessons learned from 10 years of translating research into practice*. Presentation to the 16th annual meeting of the Practice Partner Research Network (PPRNet), Medical University of South Carolina, Charleston, SC.

Chapter 24

Evaluation of Personal Health Services and Records

Morgan Price, Paule Bellwood, Ryan Habibi, Simon Diemert, Jens Weber

24.1 Introduction

Information and Communication Technology (ICT) has changed information management practices at points of care, but it is also empowering patients and individuals to take a more active role in their health and care. Through consumer-focused health ICT, such as personal health records and personal health services (e.g., health apps), people have the ability to be more engaged in their health. This is a rapidly expanding market, yet the body of evidence showing the benefits of these tools is smaller than it should be given the size of the market. Before we describe some of the evidence, we should define some of the types of consumer-focused health ICT.

24.1.1 Definitions

There are many different terms used to describe aspects of consumer health ICT with, of course, sometimes overlapping and confusing definitions. For this chapter, we will define and use the following:

- Personal Health Service(s)
- Personal Health Record
- Personal Health Information

Personal Health Services (PHS) are more broadly defined than PHRs. These are any consumer-focused health ICT tools that can help people to engage in

their own care. We have included PHRs in the broader taxonomy of PHS (see Figure 24.1). PHS do not necessarily have the mandate to provide a longitudinal record and can be focused on a specific aspect of healthcare or wellness. For example, they could provide information about foods or they could be a diet mobile health app that lets you track your diet. A PHS could support home telemedicine or it could be an activity tracker. More streamlined services have the advantage of focusing on a particular health behaviour (e.g., quitting smoking, screening for a diagnosis, or improving health literacy about a condition) and may be used in a targeted way to support a specific health issue, assess for current risk, or help a person with a behaviour change.

A **Personal Health Record** (PHR), also sometimes referred to as Personal Controlled Health Record (PCHR), is an ICT application designed to allow patients (or their designated caregivers) to store and manage their personal health information (PHI). The American Health Information Management Association (AHIMA) defines the PHR as an:

> electronic, universally available, lifelong resource of health information needed by individuals to make health decisions. Individuals own and manage the information in the PHR, which comes from healthcare providers and the individual. The PHR is maintained in a secure and private environment, with the individual determining rights of access. The PHR is medical and health information that is directed and maintained by the patient and is separate from and does not replace the legal record of any provider. (AHIMA, 2005)

The specific data elements stored within a PHR varies between different application providers. Table 24.1 has some examples.

Some PHR systems are highly comprehensive, storing a wide amount of information about patients. In other cases, the PHR application may deliberately be narrow in scope in an effort to maintain a separation between consumer information and that in the custody and control of a healthcare provider, but still maintain the concept of a longitudinal record.

PHRs that are tightly connected to a provider-based Electronic Medical Record (EMR) and represent subsets of the data represented in the corresponding provider record are called *tethered* PHRs. In contrast, *untethered* PHRs are stand-alone and may provide users with the functionality to export/import their personal health data to/from selected provider-based EMR systems, based on defined interoperability interfaces.

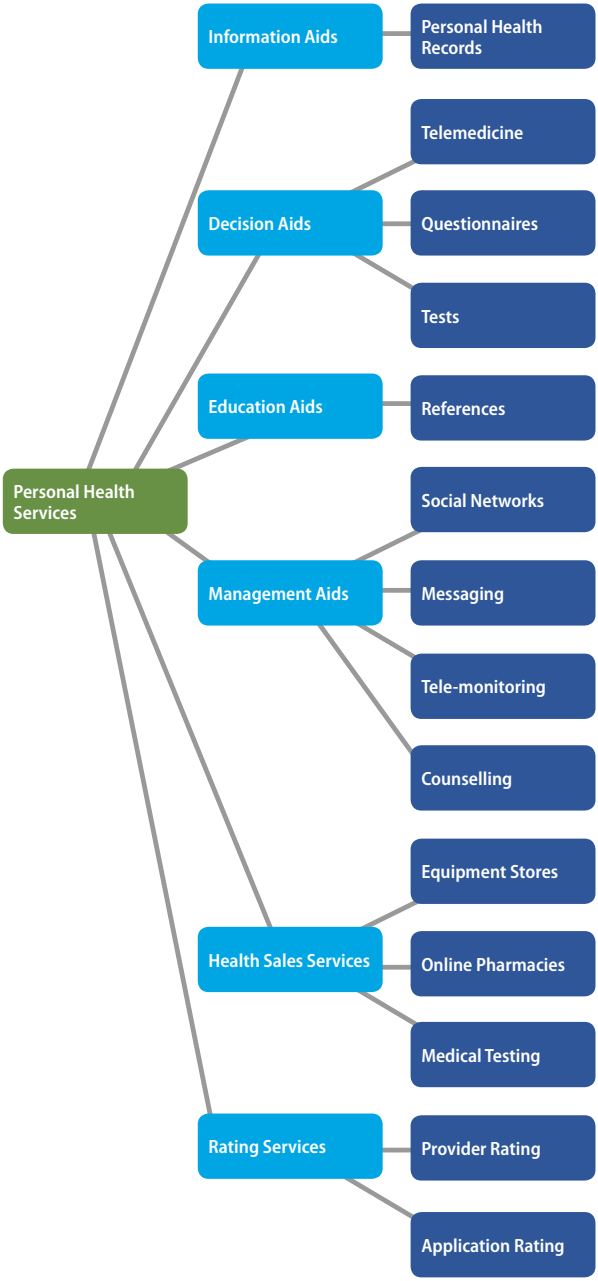


Figure 24.1. A breakdown of the broad range of personal health services.

Table 24.1
Typical Elements in a PHR (based on AHIMA, 2005)

1. Personal identifiers, such as name and date of birth
2. Emergency contact information.
3. Names, addresses, and phone numbers of physician, dentist, and other specialists
4. Health insurance information
5. Living wills and advance directives
6. Organ donor authorization
7. A list and dates of significant illnesses and surgeries
8. Current medications and dosages
9. Immunizations and their dates
10. Allergies
11. Important events, dates, and hereditary conditions in family history
12. A recent physical examination
13. Opinions of specialists
14. Important tests results
15. Eye and dental records
16. Correspondence with providers
17. Permission forms for release of information, operations, and other medical procedures
18. Any other miscellaneous information about patient health such as exercise regimen, herbal medications, and any counselling.

PHRs are a place to store and manage personal health information (defined below). Thus they can be considered an information aid for patients: a place to review, recall or share personal health information when needed to support care. However, PHRs may also integrate patient-centric knowledge bases or decision-support that extends beyond the basic function of storing information. Such advanced functionality may help with wellness activities, the management of chronic diseases or other targeted health problems, such as addictions, obesity, and mental health. Some of these functions are also available in Personal Health Services, so there is admittedly overlap between a focused PHR and robust Personal Health Services.

Personal Health Information (PHI), in contrast to both PHS and PHR, is not an application where the information resides, but is the information about an individual. It is information about an individual and that individual's health, and can include information on diagnoses, medications, encounters with care, lab results, health activities, and functional status. Table 24.1 provides application functions and also types of PHI.

Personal health information can reside in a number of ICT systems from consumer-focused ICT to provider-focused ICT and from health ICT to non-health ICT systems, such as government systems or insurance systems.

24.2 Potential Benefits of PHS

Personal Health Services have many potential benefits to multiple stakeholders. This assumes that the PHS is properly designed, implemented, promoted, adopted, and, more importantly, that it offers services that the users need and

find useful. Some of the reported potential benefits include: improving patient engagement in and accountability for their own care; enabling patients to better manage their health information and the information of their family members; providing essential information to patients and other healthcare providers in emergencies or while travelling; improving communication between patient and provider; and reducing administrative costs (Tang, Ash, Bates, Overhage, & Sands, 2006). More specifically, potential benefits can be grouped into three broad categories: (a) benefits to the consumer (i.e., the intended user of the PHS — the patient); (b) benefits to the consumer's circle of care (i.e., caregivers, healthcare providers); and (c) benefits to the overall healthcare system.

24.2.1 Potential Benefits to the Consumer

One often-stated purpose of using a PHS, such as a PHR, is supporting the user to engage in their care through accessing credible health information. This can include both personal health information (their own PHI) and general health information related to their health, such as information on medications, health conditions, or how to exercise. Consumers can use credible and evidence-based information to become better informed about their health, which allows them to improve their own illness and wellness management. Many chronic conditions require a degree of self-management, such as lifestyle changes, adherence to medications, self-monitoring (e.g., blood pressure, blood sugars). PHS can enable users to better manage their own chronic conditions by providing tools, reminders and feedback. The chronic care model (Bodenheimer, Wagner, & Grumbach, 2002) highlights the need for engaged patients, and PHS can be one way of both engaging and empowering patients in their chronic disease management.

24.2.2 Potential Benefits to the Circle of Care

PHS can improve communication between users and healthcare providers, such as enabling users to provide information on function between visits, ask more informed questions, as well as manage prescriptions, refills, and appointments (Tang et al., 2006). Further, when patients share their PHI with healthcare providers, the providers can gain valuable information on daily function, adherence, behaviours and symptoms that might not be easily captured during visits for care. This can help with decision-making, lead to improved communication, and result in better overall understanding of the issues around the progress of a disease or wellness management, both by the provider and by the patient (Tang et al., 2006). Informal caregivers, too, can benefit from access to a patient's PHR as a tool for communicating across the team, and to better understand the needs and treatments and rationale for treatments.

PHS can also provide another treatment option for providers to offer to patients. As evidence develops, providers will be able to increasingly suggest PHS options to help people with a range of health conditions such as asthma, diabetes, fertility, glaucoma, HIV, hyperlipidemia, and hypertension (Price et al., 2015) and, in all likelihood, other conditions in the future.

24.2.3 Potential Benefits to the Healthcare System

Potential PHS benefits to the healthcare system include reduced healthcare costs due to the potential improved management of various chronic conditions. This, however, is very much dependent on the actual capabilities of the PHS and how well it is adopted by its users. In addition, PHS have the potential to improve management of overall wellness; they emphasize prevention, which, in turn, may help reduce overall healthcare system costs in time (Tang et al., 2006). Although there is a potential, this is far from proven and there is much evaluation to do to better understand the impact of PHS. Also to be considered is the effort that consumers will put into managing their PHI through these various services (Ancker et al., 2015). Despite the potential benefits, the evidence for PHSS and PHRS is limited and there are challenges to adopting these tools.

24.3 Challenges for PHS

Personal Health Services, especially digital PHS, are relatively new and rapidly evolving. We do not yet know all of the positive impacts or the unintended consequences of these tools.

24.3.1 Accuracy & Safety

One challenge that has been considered is the fact that the accuracy of PHI recorded online by patients (and their informal caregivers) is dependent on the way it is collected, not to mention other factors such as computer literacy and age of the person recording the information (Kim & Kim, 2010). The provenance[1] of the PHI entered into PHR applications is important for judging its accuracy. For example, PHI data such as prescriptions and diagnoses that are entered by patients based on recalling their memories of prior visits with care providers may have lower accuracy than data directly downloaded from provider-facing (clinical) information systems or entered based on written reports. Conversely, data is that is recorded by people prospectively about their behaviours (e.g., diet, exercise, medication adherence) may be more accurate than what is recalled or described in a physician visit. PHRS have been found effective in increasing the data quality of provider medication lists (Wright et al., 2008). Provenance of PHI is increasingly important as PHS and other systems are interconnected. Unfortunately, provenance information is rarely kept in PHS and PHRS, which may compromise the objective of ensuring accuracy of PHI.

While PHRS are not considered medical devices in the classical sense, their implementation may introduce hazards that require careful consideration. Patient safety with PHS is a voiced concern from a provider perspective, both in terms of considering data of unclear accuracy and origin in clinical decision-

1 Provenance is lineage of data, such as who entered the data, who may have approved it, reviewed it, and modified it over time.

making, as well as in terms of the potential safety ramifications of allowing patients to access clinical data that they may not properly understand (Wynia, Torres, & Lemieux, 2011).

Patient controlled PHRs have been a safety concern in cases where patients are free to withhold certain information from providers and in emergency situations (Chen & Zhong, 2012) or when access or sharing is not clear. Conversely, patients may assume, incorrectly, that data is immediately shared and a message or comment that is urgent and written in a PHS or PHR is viewed by a healthcare provider, for example, when it might not be. The reverse is also true, as it has been argued that intelligent "assistant" services based on PHRs can help improve the safety of certain consumers, for example by providing self-management support to patients with heart failure (Ferguson et al., 2010).

24.3.2 Health and Technology Literacy

PHS and PHRs can provide many potential benefits but may also create new barriers, in particular for populations with low technological or health literacy. The adoption of various PHSs may create a health "digital divide". Evidence for the significant impact of technology literacy has been shown in several studies (Hilton et al., 2012; Wagner et al., 2012). Age has been validated as a predictor for technology literacy. In a randomized trial, Wagner et al. (2012) found that likelihood of PHR use decreased with age. Technology literacy in elderly populations has shown to be a significant barrier. Kim and colleagues have shown that low-income, elderly populations have a significant disadvantage of accessing online PHR services (Kim et al., 2007, 2009). These results agree with studies by Lober et al. (2006), who also researched the impact of cognitive impairment and disability in elderly populations.

Consumers do not commonly understand the medical terminology used by providers or in provider-centric records. Translating that terminology to plain language that is accessible to consumers requires significant effort if done manually. Automated solutions have been developed based on ontological engineering methods (Bonacina, Marceglia, Bertoldi, & Pinciroli, 2010) and data extraction from social health networks (Doing-Harris & Zeng-Treitler, 2011). Aside from the terminology, there is the question of how much support consumers need in documenting and interpreting important medical information, in particular their online test results. One study of consumer support needs indicated that educational and psychosocial support services were less frequently used than technical support (Wiljer et al., 2010).

24.3.3 Privacy and Security

PHI may be highly sensitive and thus needs to be carefully protected. There is significant interest in PHI from a variety of legitimate parties, including various sectors of industry (e.g., pharmaceuticals and marketers), employers, insurers, but also for fraudulent use (e.g., identity theft, credit crime). Besides patient privacy, provider privacy must also be considered, as the PHR may open up in-

formation to consumers and other parties that has traditionally been kept in private EHRs or EMRs, accessible only to physicians.

Privacy concerns are among the most important barriers perceived by both patients (Chhanabhai & Holt, 2007; Hoerbst, Kohl, Knaup, & Ammenwerth, 2010; Wen, Kreps, Zhu, & Miller, 2010) and providers (Wynia et al., 2011). Although the PHI maintained in PHS is equally sensitive to that information maintained in provider-facing systems, PHR systems are not generally subject to the same privacy regulations and legal protections.

Granular privacy controls that let consumers choose what data to share with which healthcare provider are easier to interpret by users. However, such an ability to withhold PHI raises significant care and liability issues (Cushman, Froomkin, Cava, Abril, & Goodman, 2010). Social networking features, while popular, are also challenging as consumers have difficulty correctly interpreting their privacy controls (Hartzler et al., 2011).

Cohort effects may be observed based on particular groups of consumer populations; younger consumers tend to be more willing to share their PHI (Cushman et al., 2010). Particularly vulnerable populations, such as consumers with conditions that are associated with social stigma, may require dedicated considerations, for example, people with mental health conditions (Ennis, Rose, Callard, Denis, & Wykes, 2011) and people living with HIV/AIDS (Kahn et al., 2010). Research on the latter population has indicated a high willingness to share PHI with providers and a lower willingness to share with other non-professionals (Teixeira, Gordon, Camhi, & Bakken, 2011).

Because of the patient-centric nature of PHSs and PHRs, traditional privacy consent directives such as identity-based access ("share PHI only with my doctor, Dr. X") and role-based access ("share my PHI with all doctors") are limited and fall short. The first alternative is considered too restrictive to support a continuum of collaborative care around the patient where the patient may have wished a new emergency room physician to have access to PHI in an emergency. The second alternative is considered too broad (i.e., providing little protection). Specific process-based privacy models have been developed in response to this problem (Mytilinaiou, Koufi, Malamateniou, & Vassilacopoulos, 2010). A related issue is emergency access to PHI in cases where the consumer is not able to provide consent (Chen & Zhong, 2012).

24.4 Current State of Evidence

While, there have been several reviews completed examining the expected and actual benefits of PHS, there is still a relative lack of evidence on the benefits of PHS. This is due, in part, to the rapidly changing nature of PHS and its various platforms. Smartphones and wearable technologies, for example, are radically altering platforms where various PHS apps are being developed.

Genitsaridi, Kondylakis, Koumakis, Marias, and Tsiknakis (2015) reviewed and evaluated 25 PHR systems based on four main requirements: free and open

source software requirement, Web-based system requirement, specific functionality requirements, and architectural / technical requirements. Only four (MyOscar, Indivo-X, Tolven, and OpenMRS) out of the 25 PHR systems reviewed met the free and open source software and Web-based requirements, which were considered as basic requirements for a PHR system regardless of its functionality level. These four PHR systems, in addition to six other highly popular PHR systems, were then evaluated based on specific functionality requirements (i.e., recording of a problem, diagnosis, and treatment, self-health monitoring, communication management, security and access control, and intelligence factors) as well as architectural requirements (i.e., stand-alone, tethered, or interconnected). This study determined that there is a need for better design of PHRs in order to improve self-management and integration into care processes (Genitsaridi et al., 2015).

There is early evidence to support the use of PHRs in some chronic conditions. Based on a systematic review, there is evidence that PHRs can be used to benefit the following: asthma, diabetes, fertility, glaucoma, HIV, hyperlipidemia, and hypertension (Price et al., 2015). There is a small body of empirical evidence demonstrating benefit; however, many of these are short-term studies looking only at changes in behaviour or early clinical outcomes.

There are many factors that can impact the realization of benefits of PHS and PHR, beyond just the features and qualities (such as usability) of the tools themselves. Thus, it is important to consider a wide range of factors in evaluation including, among others: the PHS tool itself; the people who use it directly; the people who use it indirectly (e.g., care providers who see summary information); the context of use; the integration with care; the incentives (e.g., incentives from health insurance). One key issue to consider when evaluating PHS is the interest and capacity of people to manage their health through electronic means. As discussed previously, a health digital divide is possible if services are available through PHS. Consider predictors of use of your users that include education, technical knowledge, and health knowledge (Kim & Abner, 2016). Thus, evaluation (and implementation training) should carefully consider the level of health and technological literacy of the users.

24.5 Selected Case Study Examples

24.5.1 Case study 1 – Kaiser Permanente's *My Health Manager*

Kaiser Permanente, one of the largest health delivery organizations in the United States, began implementing PHR solutions for their members in 2004. The PHR platform, *My Health Manager*, was tethered to their electronic health record (EHR) and included not only information services, but also provided means for secure communication between patients and providers. The system was well received and had been adopted by 2.4 million patients by 2008 (Silvestre, Sue, & Allen, 2009). By 2013, 65% of all eligible Kaiser Permanente

members were registered in *My Health Manager*. Early studies showed a significant decrease in office visits (26.2%) within a period of three years, while at the same time there was a ninefold increase in online consultations (phone visits) and a dramatic increase in patient-generated secure messages (Chen, Garrido, Chock, Okawa, & Liang, 2009). Member satisfaction and health outcomes remained largely unchanged over the three-year study, with a few exceptions, particularly with respect to certain chronic disease conditions such as HbA1c control, antidepressant medication management, and osteoporosis management in female populations, which developed negatively. Further studies have also shown that the PHR use has been correlated with significant health benefits in subpopulations such as people with diverse languages and ethnicity (Garrido et al., 2015). However, language and ethnicity both influenced the likelihood of members signing up to the PHR system.

A recent study on Kaiser Permanente's patient outcome improvements focused on virtual doctor-patient communication (Reed, Graetz, Gordon, & Fung, 2015). *My Health Manager* provides the ability for patients and providers to communicate over e-mail as well as schedule appointments and maintain many other health management aspects online. Over 50% of study participants had used the e-mail feature at least once, and almost 50% of participants prefer e-mail as the first method of contact when it comes to their medical concerns. This resulted in 42% of respondents reporting a reduction in phone contact and 36% of respondents reporting a reduction in in-person visits. Overall, the use of the *My Health Manager* system resulted in 32% of users with chronic conditions improving their overall health (Reed et al., 2015). In addition, the results of another study suggest that using tools for health care management (i.e., online medication refills) can result in improving medication adherence (Lyles et al., 2016).

Kaiser Permanente's portal also provides users with access to information about prevention, health promotion, and care gaps. In addition to improved communication and reduction in office visits and phone calls, users of *My Health Manager* are more likely to participate in certain preventive measures, such as cancer screening, hemoglobin A1c testing, and pneumonia vaccination (Henry, Shen, Ahuja, Gould, & Kanter, 2016).

24.5.2 Case study 2 – English National Health Service's *HealthSpace*

The National Health Service (NHS) in England attempted an implementation of a public nationwide PHR called *HealthSpace* in 2007. A three-year evaluation was completed by the Healthcare Innovation and Policy Unit at the London School of Medicine and Dentistry (Greenhalgh, Hinder, Stramer, Bratan, & Russell, 2010). It was initially inspired by the Kaiser Permanente model outlined above. The NHS' goals for this PHR were personalizing care, empowering patients, reducing NHS costs, and improving data quality and health literacy. *HealthSpace* included a basic account that would allow a person to record their own data (e.g., blood pressures) and an advanced account where they could gain access to their summary care record (a subset of PHI shared from the patient's

GP) and interact with their GP (to book appointments, message with questions). Additional features were planned over time.

The evaluation of *HealthSpace* was a mixed method, multilevel case study. It covered the policy development, implementation, and patient experience using both qualitative and quantitative methods to develop a rich picture of *HealthSpace*.

The policy and project documentation that was evaluated in this case study highlighted a focus on the technical and managerial aspects of implementing a PHR, with less focus on understanding the user requirements (e.g., through observation and detailed analysis and testing). The evaluation highlighted a design gap in user expectations and needs with respect to how the system was implemented. The deployment of this particular PHR, unfortunately, resulted in poor initial uptake mostly due to a lack of interest, perceived usefulness and ease of use, and a cumbersome account creation process. During the PHR evaluation, *HealthSpace* users expressed disappointment in specific data being unavailable, the need for data self-entry, and an inability to share their information with their healthcare providers seamlessly. The study highlighted that *HealthSpace* was not aligned with the "attitudes, self-management practices, [and] identified information needs" of its potential users (Greenhalgh et al., 2010). The expected benefits of *HealthSpace* were not realized, in large part, due to this gap.

24.6 Issues, Guidance and Implications

PHS and PHRs have the potential for wide ranging impact on care — both directly for the patient and indirectly for the care providers, care organizations, and the overall healthcare system. Thus, we suggest considering evaluation using a broad framework such as the Clinical Adoption Framework (see chapter 3), which includes concepts from micro-level evaluation (system, use, and patient level outcomes) to meso-level and macro-level influencing factors. Also, we encourage the use of multiple methods when evaluating PHS, and a plan that incorporates various assessments to occur over time to see how the PHSs are incorporated into health and wellness behaviours and into healthcare systems. With multi-method studies, one can also develop feedback loops into the PHS programs, using evaluation in an action research framework to improve the chance of success and positive impact of using these tools. Large, single trials, at this stage, may not be able to provide the richness of answers needed to understand how PHSs are being used and why they are achieving (or not achieving) their outcomes. Also, it is important to consider how to incorporate the rate of change of PHS features and functions into the evaluation, as these are rapidly evolving tools.

For example, evaluation can begin prior to system implementation by modelling out the goals of the PHS implementation and related activities and mapping these into the meso- and macro-level contexts. This may, for example, quickly highlight disconnects between the goals of the PHS and macro-level aspects such

as legislation or funding limitations for providers (e.g., no mechanism for remuneration for e-communication). Usability evaluations (both usability inspections with experts and usability testing with potential users) can be completed with early prototypes. Once implemented, pilot studies can explore user experience as well as the indirect experience of providers when patients have access to PHS. Future studies can then begin to look at changes in behaviour and changes in outcomes, both clinical and health system (e.g., numbers of visits, numbers of e-visits, and capacity to see patients).

24.7 Summary

PHSs and PHRs are being increasingly implemented as part of health care systems. Despite the efforts in implementation and adoption, the advertising of apps and wearables, et cetera, there is still a gap in sufficient evaluation of PHS. We need a better understanding of how these tools are used and what the impact these tools have on long-term outcomes, both health outcomes and such health system outcomes as capacity and cost.

When planning an evaluation for PHS it is important to consider the goals and plan an evaluation based on those goals and the potential direct and indirect impacts over time. Unintended consequences should be considered. Depending on the scope of the PHS, the evaluation should be broad, assessing impact across the continuum of care (i.e., across the patient's circle of care). To do this, we advocate for multi-method studies that will evaluate the design and adoption of the PHS tools early and throughout its life cycle. A deeper understanding of user needs early (e.g., during concept design, the establishment of projects, the development of policy) will better ensure that the final product meets the actual needs of users. Finally, consider evaluation across the range of dimensions in the Clinical Adoption Framework (see chapter 3) to provide a breadth that is needed to understand the impact of PHS across the micro, meso and macro levels of the healthcare system.

References

American Health Information Management Association [AHIMA]. (2005). The role of the personal health record in the EHR. *Journal of AHIMA, 76*(7), 64A–64D.

Ancker, J. S., Witteman, H. O., Hafeez, B., Provencher, T., Van de Graaf, M., & Wei. E. (2015). The invisible work of personal health information management among people with multiple chronic conditions: Qualitative interview study among patients and providers. *Journal of Medical Internet Research, 17*(6), e137. doi: 10.2196/jmir.4381 PMID: 26043709 PMCID: 4526906

Bodenheimer, T., Wagner, E. H., & Grumbach, K. (2002). Improving primary care for patients with chronic illness. *Journal of the American Medical Association, 288*(14), 1775–1779.

Bonacina, S., Marceglia, S., Bertoldi, M., & Pinciroli, F. (2010). Modelling, designing, and implementing a family-based health record prototype. *Computers in Biology and Medicine, 40*(6), 580–590. doi: 10.1016/j.compbiomed.2010.04.002

Chen, C., Garrido, T., Chock, D., Okawa, G., & Liang, L. (2009). The Kaiser Permanente electronic health record: Transforming and streamlining modalities of care. *Health Affairs, 28*(2), 323–333. doi: 10.1377/hlthaff.28.2.323

Chen, T., & Zhong, S. (2012). Emergency access authorization for personally controlled online health care data. *Journal of Medical Systems, 36*(1), 291–300. doi: 10.1007/s10916-010-9475-2

Chhanabhai, P., & Holt, A. (2007). Consumers are ready to accept the transition to online and electronic records if they can be assured of the security measures. *Medscape General Medicine, 9*(1), 8.

Cushman, R., Froomkin, A. M., Cava, A., Abril, P., & Goodman, K. W. (2010). Ethical, legal and social issues for personal health records and applications. *Journal of Biomedical Informatics, 43*(5), S51–S55. doi: 10.1016/j.jbi.2010.05.003

Doing-Harris, K., & Zeng-Treitler, Q. (2011) Computer assisted update of consumer health vocabulary through mining of social network data. *Journal of Medical Internet Research, 13*(2), e37. doi: 10.2196/jmir.1636

Ennis, L., Rose, D., Callard, F., Denis, M., & Wykes, T. (2011). Rapid progress or lengthy process? Electronic personal health records in mental health. *BioMed Central Psychiatry, 11*(1), 117. doi: 10.1186/1471-244X-11-117

Ferguson, G., Quinn, J., Horwitz, C., Swift, M., Allen, J., & Galescu, L. (2010). Towards a personal health management assistant. *Journal of Biomedical Informatics, 43*(5 Suppl), S13–S16. doi: 10.1016/j.jbi.2010.05.014

Garrido, T., Kanter, M., Meng, D., Turley, M., Wang, J., Sue, V., & Scott, L. (2015). Race/ethnicity, personal health record access, and quality of care. *American Journal of Managed Care, 21*(2), e103–e113.

Genitsaridi, I., Kondylakis, H., Koumakis, L., Marias, K., & Tsiknakis, M. (2015). Evaluation of personal health record systems through the lenses of EC research projects. *Computers in Biology and Medicine, 59*, 175–185. doi: 10.1016/j.compbiomed.2013.11.004

Greenhalgh, T., Hinder, S., Stramer, K., Bratan, T., & Russell, J. (2010). Adoption, non-adoption, and abandonment of a personal electronic health record: Case study of HealthSpace. *BMJ: British Medical Journal, 341*(7782), 1091. doi: 10.1136/bmj.c5814

Hartzler, A., Skeels, M. M., Mukai, M., Powell, C., Klasnja, P., & Pratt, W. (2011, October). *Sharing is caring, but not error free: transparency of granular controls for sharing personal health information in social networks*. Proceedings of the American Medical Informatics Association Annual Symposium, Washington, DC (pp. 559–568). Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3243199&tool=pmcentrez&rendertype=abstract

Henry, S. L., Shen, E., Ahuja, A., Gould, M. K., & Kanter, M. H. (2016). The online personal action plan. *American Journal of Preventive Medicine, 51*(1), 71–77. doi: 10.1016/j.amepre.2015.11.014

Hilton, J. F., Barkoff, L., Chang, O., Halperin, L., Ratanawongsa, N., Sarkar, U., … Kahn, J. S. (2012). A cross-sectional study of barriers to personal health record use among patients attending a safety-net clinic. *PloS One, 7*(2), e31888. doi: 10.1371/journal.pone.0031888

Hoerbst, A., Kohl, C. D., Knaup, P., & Ammenwerth, E. (2010). Attitudes and behaviors related to the introduction of electronic health records among Austrian and German citizens. *International Journal of Medical Informatics, 79*(2), 81–89. doi: 10.1016/j.ijmedinf.2009.11.002

Kahn, J. S., Hilton, J. F., Van Nunnery, T., Leasure, S., Bryant, K. M., Hare, C. B., & Thom, D. H. (2010). Personal health records in a public hospital: experience at the HIV/AIDS clinic at San Francisco General Hospital. *Journal of the American Medical Informatics Association, 17*(2), 224–228. doi: 10.1136/jamia.2009.000315

Kim, S., & Abner, E. (2016). Predictors affecting personal health information management skills. *Informatics for Health & Social Care, 41*(3), 211.

Kim, E. -H., & Kim, Y. (2010). *Digital divide: Use of electronic personal health record by different population groups*. Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Buenos Aires (pp. 1759–1762). doi: 10.1109/IEMBS.2010.5626732

Kim, E. -H., Stolyar, A., Lober, W. B., Herbaugh, A. L., Shinstrom, S. E., Zierler, B. K., … Kim, Y. (2007). *Usage patterns of a personal health record by elderly and disabled users*. American Medical Informatics Association Annual Symposium Proceedings, Chicago (pp. 409–413). Retrieved from http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2655817&tool=pmcentrez&rendertype=abstract

Kim, E. –H., Stolyar, A., Lober, W. B., Herbaugh, A. L., Shinstrom, S. E., Zierler, B. K., … Kim, Y. (2009). Challenges to using an electronic personal health record by a low-income elderly population. *Journal of Medical Internet Research, 11*(4), e44.

Lober, W. B., Zierler, B., Herbaugh, A., Shinstrom, S. E., Stolyar, A., Kim, E. H., & Kim, Y. (2006). *Barriers to the use of a personal health record by an elderly population*. American Medical Informatics Association (AMIA) Annual Symposium Proceedings, Washington, DC (pp. 514–518).

Lyles, C. R., Sarkar, U., Schillinger, D., Ralston, J. D., Allen, J. Y., Nguyen, R., & Karter, A. J. (2016). Refilling medications through an online patient portal: Consistent improvements in adherence across racial/ethnic groups. *Journal of the American Medical Informatics Association, 23*(e1), e28–e33. doi: 10.1093/jamia/ocv126

Mytilinaiou, E., Koufi, V., Malamateniou, F., & Vassilacopoulos, G. (2010). A context-aware approach to process-based PHR system security. *Studies in Health Technology and Informatics, 156*, 201–213.

Price, M., Bellwood, P., Kitson, N., Davies, I., Weber, J., & Lau, F. (2015). Conditions potentially sensitive to a personal health record (PHR) intervention, a systematic review. *BioMed Central Medical Informatics and Decision Making, 15*(1), 32. doi: 10.1186/s12911-015-0159-1

Reed, M., Graetz, I., Gordon, N., & Fung, V. (2015). Patient-initiated e-mails to providers: Associations with out-of-pocket visit costs, and impact on care-seeking and health. *The American Journal of Managed Care, 21*(12), e632.

Silvestre, A., Sue, V. M., & Allen, J. Y. (2009). If you build it, will they come? The Kaiser Permanente model of online health care. *Health Affairs, 28*(2), 334–344. doi: 10.1377/hlthaff.28.2.334

Tang, P. C., Ash, J. S., Bates, D. W., Overhage, J. M., & Sands, D. Z. (2006). Personal health records: Definitions, benefits, and strategies for overcoming barriers to adoption. *Journal of the American Medical Informatics Association, 13*(2), 121–126. doi: 10.1197/jamia.M2025

Teixeira, P. A., Gordon, P., Camhi, E., & Bakken, S. (2011). HIV patients' willingness to share personal health information electronically. *Patient Education and Counseling, 84*(2), e9–e12. doi: 10.1016/j.pec.2010.07.013

Wagner, P. J., Howard, S. M., Bentley, D. R., Seol, Y., & Sodomka, P. (2010). Incorporating patient perspectives into the personal health record: Implications for care and caring. *Perspectives in Health Information Management / AHIMA, American Health Information Management Association, 7*, 1e.

Wagner, P. J., Dias, J., Howard, S., Kintziger, K. W., Hudson, M. F., Seol, Y., & Sodomka, P. (2012). Personal health records and hypertension control: A randomized trial. *Journal of the American Medical Informatics Association, 19*(4), 626–634. doi: 10.1136/amiajnl-2011-000349

Wen, K., Kreps, G., Zhu, F., & Miller, S. (2010). Consumers' perceptions about and use of the internet for personal health records and health information exchange: Analysis of the 2007 health information national trends survey. *Journal of Medical Internet Research, 12*(4), e73. doi: 10.2196/jmir.1668

Wiljer, D., Leonard, K. J., Urowitz, S., Apatu, E., Massey, C., Quartey, N. K., & Catton, P. (2010). The anxious wait: Assessing the impact of patient accessible EHRs for breast cancer patients. *BioMed Central Medical Informatics and Decision Making, 10*(1), 46–46. doi: 10.1186/1472-6947-10-46

Wright, A., Poon, E. G., Wald, J., Schnipper, J. L., Grant, R., Gandhi, T. K., … Middleton, B. (2008). *Effectiveness of health maintenance reminders provided directly to patients.* Proceedings of the American Medical Informatics Association (AMIA) Annual Symposium, Washington, DC (p. 1183). Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18999087

Wynia, M. K., Torres, G. W., & Lemieux, J. (2011). Many physicians are willing to use patients' electronic personal health records, but doctors differ by location, gender, and practice. *Health Affairs (Project Hope), 30*(2), 266–273. doi: 10.1377/hlthaff.2010.0342

Chapter 25

Evaluating Telehealth Interventions

Anthony J. Maeder, Laurence S. Wilson

25.1 Introduction

This chapter discusses an area viewed by many as a "special case" in eHealth evaluations: dealing with usage of *telehealth*, which is the delivery of healthcare services of a clinical nature where the provider of the service is remote in location and/or time from the recipient (such as teleconsultation, or teleradiology). We use the term *telehealth intervention* to indicate that our focus is on clinical processes (such as diagnosis or therapy) employing telehealth as a major component of their delivery. This term implies that the telehealth aspect is overlaid or inserted in a broader clinical activity or service, of which other components may be achieved by non-telehealth means.

Within the scope of our discussion, we also include evaluation of projects that establish and deploy these types of interventions, but not the evaluation of health services or systems as a whole, within which the interventions are delivered as one of a set of diverse and often complex interconnected components. This exclusion applies also to regional and national telehealth systems which serve multiple purposes and are therefore in the domain of health enterprise evaluation, rather than directly tractable by analysis methods intended for clinical services. An approach to such broader analysis is exemplified by work undertaken in Canada to develop a set of National Telehealth Outcome Indicators (Scott et al., 2007), which provided a base set of measurable indicators in the areas of quality, access, acceptability and costs, for post-implementation service-based evaluations. We also exclude the evaluation of underlying ICT-based mechanisms and infrastructure, including networks and systems that transmit and support telehealth such as broadband communications connectivity, and turnkey videoconferencing or store-and-forward systems, which are able to be

suitably evaluated by application of established technology or information systems analysis methods.

In the following sections we will first discuss how perspectives on telehealth can impact philosophically on evaluation approaches, imposing in some cases limitations and a narrowed view, which can discourage inclusion of a "full spectrum" of potential elements in evaluations. We will identify a wide range of approaches and associated elements that may be considered appropriate for telehealth evaluations, drawing predominantly from contributions in the clinical literature. Next we will link these elements with frameworks for evaluation that have been suggested by several authors, to demonstrate that the same elements may be viewed in different combinations and targeting different evaluation purposes. Finally, we will provide a commentary on practical constraints and considerations when conducting telehealth evaluations, and illustrate this with a case study based on a stand-alone intervention project.

25.2 Background

Early work in telehealth was poorly served by inadequate evaluation efforts. There are several reasons for this deficiency. Emphasis was often placed on the novelty of the technology or organizational aspects of the intervention, leading to evaluation of these aspects in preference to others more relevant to health impacts, and using associated evaluation methods which were often unfamiliar in clinical settings. A simplistic initial view of telehealth as the utilization of one of only a few different IT delivery mechanisms (such as video or image transfer), which could be analysed separately from any human or organizational aspects, reinforced this viewpoint. Health benefits and health economics gains are typically realized only after a lengthy period of time, beyond the extent of projects which delivered the intervention. Consequently, long-term clinical quality of care improvements and health services efficiency gains have often been regarded as impractical to evaluate. On the other hand, participant experience and satisfaction is relatively easy to assess, and so many early evaluations incorporated that as a significant component, a trend that has continued.

As noted by Bashshur, Shannon, and Sapci (2005), a dilemma exists as to whether to evaluate a telehealth intervention as if it were a typical health intervention coincidentally delivered by telehealth technology, or whether to treat it as a special type of intervention for the purpose of evaluation, because it relies on telehealth. A related issue arising is whether conventional evaluation methods for health interventions generally are applicable to telehealth interventions, as the first model above would imply, or whether specific evaluation methods should be developed for telehealth, in line with the second model. In reality, telehealth interventions are seldom evaluated without substantial interest in the telehealth aspects, so the second model has tended to dominate evaluation approaches. Consequently, evaluation methods designed for eHealth such as STARE-HI and GEP-HI in the clinical process arena, or for technology-based

health interventions more generally such as TAM and UTAUT in the user arena, are often deemed inadequate for telehealth interventions.

25.3 Telehealth Evaluation Approaches

Initial formal contributions in the field proposed flexible approaches concentrating on case-specific aspects of interest (Bashshur, 1995) or selective use of generic health services measures. For example, Hailey, Jacobs, Simpson, and Doze (1999) proposed that evaluation be performed across five areas: *specification, performance measures, outcomes, summary measures,* and *operational considerations*. *Cost* and *workload* aspects were identified as an important specific area, warranting careful development of appropriate analysis methods (Wootton & Hebert, 2001), and these have subsequently been a focus of many studies. Another important area targeted by many researchers was *psychosocial aspects* related to users (Stamm, Hudnall, & Perednia, 2000), such as usability and satisfaction. Emphasis was also placed on the efficacy of *diagnostic and management decisions* (Hersch et al., 2002) and associated impacts on *access* and *outcomes* in telehealth services (Hersch et al., 2006). Furthermore, *technical* aspects of implementations were also seen as a part of evaluation (Clarke & Thiyagarajan, 2008), in the areas of *information capture and display*, and *information transmission* (including statistical analysis and visual quality).

The notion of inferred *causality* linking the intervention characteristics with observed effects which were ascribed to telehealth in evaluations was described by Bashshur et al. (2005), and the influence of *medical care process models* for unifying the effects of client and provider behaviours and explaining participation effects and clinical outcomes was advocated by Heinzelmann, Williams, Lugn, and Kvedar (2005). These two alignments suggest that one strategy for conducting evaluations is to focus predominantly on the clinical aspects, which Brear (2006) has typified as determining *clinical benefits*, causal *influences* from *technical, people* and *organizational* factors, and *cost-effectiveness* in terms of obtaining the benefits (see Figure 25.1 below).

Alternatively, approaches to evaluation can be derived through synthesis, by identifying key groupings of evaluation elements from reviews of studies of a number of comparable interventions. Ekeland, Bowes, and Flottorp (2010) reviewed a wide range of studies offering evidence of clinical effectiveness and itemized major evaluation elements as *behavioural, cost/economic, health, organizational, perception/satisfaction, quality of life, safety, social,* and *technology*. Deshpande and colleagues (2009) reviewed store-and-forward interventions and summarized the main evaluation elements in four categories: *health outcomes, process of care, resource utilization* and *user satisfaction*. Wade, Kanon, Elshaug, and Hiller (2010) considered economic analyses of telehealth services, and determined that evaluation elements could be grouped as *costs and effects, technology*, and *organizational aspects*.

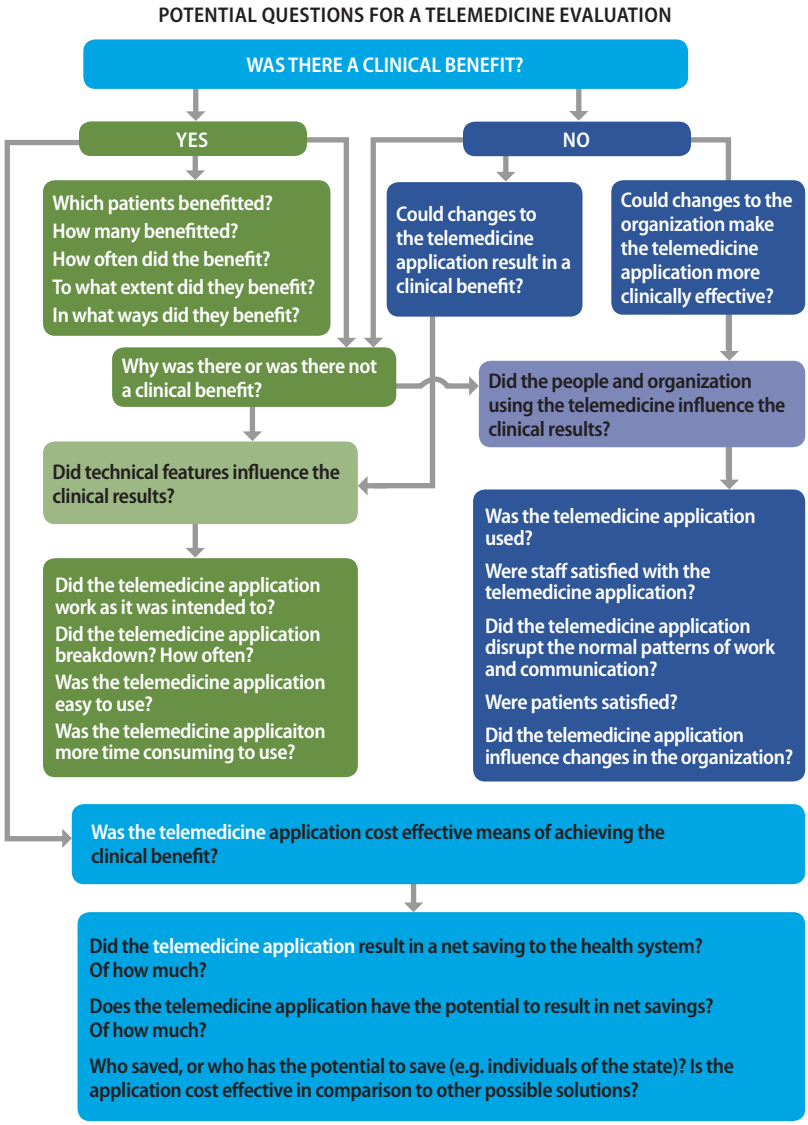


Figure 25.1. Clinically focused evaluation strategy.

Note. From "Evaluating telemedicine: lessons and challenges," by M. Brear, 2006, *The Health Information Management Journal* (Australia), *35*(2), p. 25. Copyright 2006 by SAGE Publications, Ltd. Reprinted with permission.

Recently a collaborative European proposal has been developed for a comprehensive Model for Assessment of Telemedicine Applications (MAST) (Kidholm

et al., 2012) which provides a wide scope of synthesis by addressing seven distinctive evaluation domains: *health problem and application, safety, clinical effectiveness, patient perspectives, economic approach, organizational aspects*, and *socio-cultural/ethical/legal aspects*. It is recommended that these be analysed in a three-step approach, covering *preceding considerations, multidisciplinary assessment*, and *transferability assessment*. This possibly is the most extensive example of a synthesis approach and has yet to see widespread adoption.

25.3.1 Telehealth Evaluation Frameworks

Evaluation frameworks have been developed to provide a higher-level contextual setting for selection, or aggregation, of the above diverse elements. An evaluation framework consists of categories containing different evaluation questions or objectives, from which an evaluator might choose those most pertinent to the intervention. A strong argument in favour of framework approaches is that ad hoc choices of evaluation elements can lead to selection (or, alternatively, omission) of measures which are strongly correlated with the success (or failure) of interventions (Jackson & McClean, 2012).

Some early framework concepts followed a sequential set of considerations related to the telehealth intervention: Hebert (2001) proposed three areas of focus for evaluation: *structure*, *process* and *outcomes*. Bashshur et al. (2005) advocated a refined version of this approach with high level sequential structuring of evaluation aspects in four time steps: *evaluability assessment* to identify what could or could not be evaluated based on the description and scope of the intervention project; *documentation evaluation* (including artefacts such as software) for the intervention design and implementation; then applying *formative or process evaluation* for the change and acceptance associated with deployment of the intervention in a clinical service; and finally *summative or outcome evaluation* applicable to health and economic benefits.

Taxonomies of telehealth are useful for identifying and grouping elements, which may be candidates for evaluation, in different circumstances. Tulu, Chatterjee, and Maheshwari (2007) defined a structural taxonomy based on the components that must be used in the realization of a service, namely *application purpose, application area, environmental setting, communication infrastructure*, and *delivery options*. More recently, Bashshur, Shannon, Krupinski, and Grigsby (2011) advanced a more top-down approach via conceptualization as a three dimensional space describing intersection sets of *functionality, application* and *technology* elements (see Figure 25.2). Nepal, Li, Jang-Jaccard, and Alem (2014) proposed a framework of broader coverage, including six aspects for evaluation: *health domains, health services, delivery technologies, communication infrastructure, environment setting*, and *socio-economic analysis*.

Alternative approaches to evaluation frameworks have emerged recently in an attempt to provide greater inclusivity and flexibility, as those described above tend to focus on abstract concepts to define them. Van Dyk (2014) reviewed possible areas for evaluation based on technology development models, and

proposed a multi-dimensional space associated with *technology maturity* principles and *systems life cycle* concepts. A hybrid approach was proposed by Maeder, Gray, Borda, Poultney, and Basilakis (2015) as a means of aligning evaluation with *organizational learning models* and *health system performance indicators*. Such frameworks as these offer comprehensive coverage and useful mechanisms for describing evaluation instances (especially those pertinent to large-scale projects or services), but add conceptual complexity that cannot be easily navigated for simpler telehealth implementations.

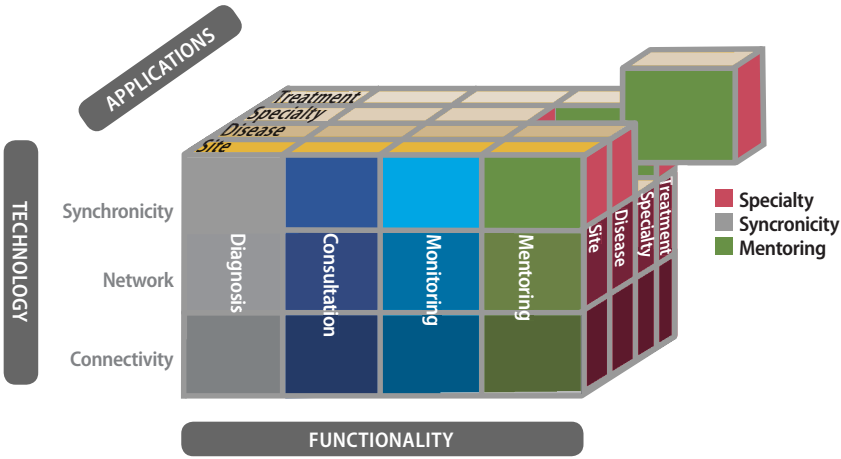


Figure 25.2. Top-down taxonomy.

Note. From "The taxonomy of telemedicine," by R. Bashshur, G. Shannon, E. Krupinski, and J. Grigsby, 2011, *Telemedicine and e-Health, 17*(6), p. 491. Copyright 2011 by Mary Ann Liebert, Inc. Publishers.

25.3.2 Telehealth Evaluation Practice

The lack of consensus on evaluation methodologies for telehealth is largely a consequence of the complexity of telehealth interventions. Many of the frameworks discussed so far represent attempts to map this complexity onto evaluation methodologies, whose aim is to measure the impact and efficacy of a telehealth intervention. The "gold standard" in the evaluation of medical interventions is the randomized controlled trial (RCT), which tends to be applied to an intervention as a self-standing analysis, without catering for the effects of contextual complexity.

There are many reasons why such a trial is not usually feasible in telehealth (Agboola, Hale, Masters, Kvedar, & Jethwani, 2014), including the inability to conceal from participants the assignment of subjects into control or intervention groups. The complexity and expense of RCTs limits their application to small, short-term projects. There is also an ethical issue of denying control groups access to apparently beneficial technologies, when the aim of the evaluation might be to assess the cost-effectiveness of an intervention whose clinical benefit might not be in dispute (Bonell, Fletcher, Morton, Lorenc, & Moore,

2012). Furthermore, there is a need in telehealth evaluations to investigate not only the change in clinical outcomes, but also the mechanisms underlying such changes. Such mechanisms should ideally be studied individually, as well as through their combined impacts on clinical outcomes. RCTs are not capable of such things as assessing the separate effects of intervention components or of discovering hidden explanations for the success or otherwise of interventions (Marchal et al., 2013).

A major telehealth evaluation exercise using cluster randomized trial methodology was conducted as part of the United Kingdom-based Whole Systems Demonstrator (WSD) project, seeking to validate the effects of home telecare on a range of clinical aspects including mortality, hospital admissions, use of care, quality of life, etc. (Steventon, Bardsley, & Billings, 2012). This provides a good example of the pros and cons of the randomized trial approach. While a high strength of evidence was obtained by sample sizes in the range of thousands, many of the findings did not show major gains for telehealth and it has been suggested that such large-scale trials may be subject to systematic bias due to their health system context (Greenhalgh, 2012).

A feature of RCTs is the separation of experimenters and participants; a double-blind trial is administered by clinicians who are unaware of which group (control or intervention) subjects belong to. As pointed out above, such methodologies produce rigorous verifiable measures, but might not capture the benefits and mechanisms of complex medical interventions such as telehealth. A growing trend is to reduce the isolation of researchers and subjects, with benefits to both assessing the benefits of interventions, and to more widespread implementation of such interventions. For example, in a wide-ranging review of participatory research by Jagosh and colleagues (2012), it was concluded that "multi-stakeholder co-governance can be beneficial to research contexts, processes, and outcomes in both intended and unintended ways".

It is clear from the preceding that telehealth is among the more complex medical interventions and, accordingly, evaluation of telehealth systems cannot adopt methodologies that might be appropriate for, say, a pharmaceutical trial. Increasingly, telehealth projects are assessed by methods in which a large number of stakeholders contribute to the process, and the underlying research questions go beyond simple measures of clinical effectiveness. It has been noted (Gagnon & Scott, 2005) that telehealth evaluation often serves different purposes for different stakeholders, so it might be expected that no single evaluation framework or methodology can cater comprehensively for it.

This complex environment may be best approached by a participatory strategy for evaluation, involving stakeholders in study designs. Translation of evaluation findings and evidence to influence policy is a further challenge, as policy-makers are typically difficult to engage as stakeholders in long-term studies; nevertheless, the power of case studies to connect back to them has been demonstrated (e.g., Jennett et al., 2004). The question of responsiveness and insight by policy-makers in response to the provision of evaluation findings and

evidence has been raised (Doarn et al., 2014) and it is argued that policy formulation might be included as a stage of any overall evaluation.

25.4 Case Study: Evaluation Using Participatory Principles

Chang (2015) identified five stages in the cycle of telehealth implementation: *inputs, activities, outputs, outcomes* and *impact*. However, in practical telehealth implementations, the early stages of the project (system design, stakeholder analysis) are often separated from other processes, mainly through such restraints as the need to use off-the-shelf hardware, or interoperability issues outside the scope of the project, or the difficulty of involving all stakeholders in the study. In cases where participants are able to contribute to technology design, such participatory methods have been shown to contribute to the success of telehealth systems (Li et al., 2006).

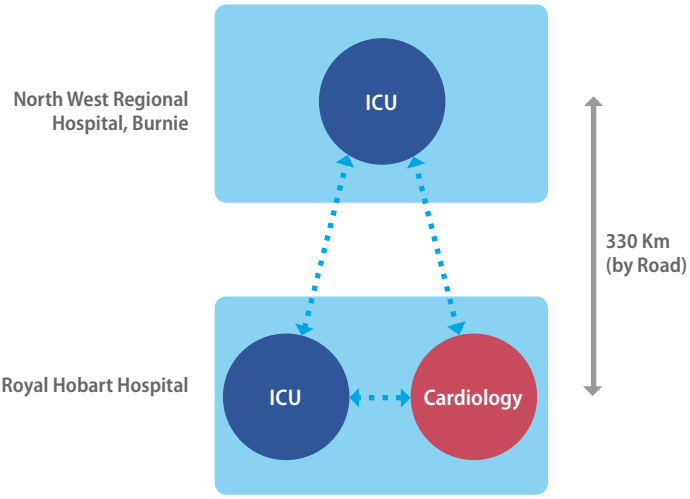


Figure 25.3. Telehealth connectivity for the case study project.

Note. From "Applying an integrated approach to the design, implementation and evaluation of a telemedicine system," by S. Hansen, L. Wilson, and T. Robertson, 2013, *Journal of the International Society for Telemedicine and eHealth, 1*(1), p. 21. Copyright 2013 by ISFTEH. CC BY License.

An example of a telehealth implementation, which incorporates aspects of participatory design and participatory research/evaluation, was the ECHONET project in Australia described by Hansen, Wilson, and Robertson (2013). Its principal aim was to support the Intensive Care Unit of North West Regional Hospital (NWRH) located in Burnie, North Western Tasmania. This ICU had basic intensivist coverage, but relied on other hospitals, and predominantly a major tertiary hospital Royal Hobart Hospital (RHH), for support in other specialist services, notably bedside echocardiography (see Figure 25.3). In this project, three mobile

multichannel broadband telemedicine units connected, over a broadband network, the ICU of NWRH with separate nodes in two departments (Cardiology and ICU) of RHH. The aim was not to provide a fully outsourced intensivist service, the suggested model for some recent eICU implementations (Goran, 2012), but to provide support for the small, isolated specialist staff at NWRH.

A combination of a participatory research philosophy and learnings from the team's previous experience with telemedicine systems (Wilson, Stevenson, & Cregan, 2009) influenced the approach. It was agreed from the beginning that an integrated design, implementation and evaluation approach would be adopted. Underpinning the practice of participatory research is an intention of the researcher to effect positive change on the situation within which the research is taking place while simultaneously conducting research, and a collaborative approach between the researcher and subject in reaching this objective and developing understanding.

Activities were carried out in the ECHONET project that informed the design of the system, the implementation strategy adopted, and the criteria assessed in the evaluation. These activities consisted of *stakeholder interviews*, *baseline study*, *design workshops,* and activities relating directly to the clinical trial of ECHONET including *interviews, questionnaires* and *logbooks*. In detail, these activities were as follows:

- The stakeholder interviews helped to establish the success criteria by which the system was assessed in the evaluation phase. They also served to inform the design workshops by establishing potential applications outside the design brief.

- The baseline study provided a datum on which changes might be captured as a result of the implementation and provided the project team with an understanding of the context and environment in which ECHONET would be used, including clinicians' existing work practices.

- Several design workshops were carried out with mock-ups of the graphical user interface (GUI) and as early prototypes became available, enabling the project to capture the benefits of user-centred design as described by Sutcliffe et al. (2010).

- Instruments deployed during the trial included weekly interviews with all users, logbooks, and a series of mid-trial interviews to monitor the trial for possible modifications, and to refine the end-of-trial processes. Post-trial instruments consisted of interviews with participants, a questionnaire for all participants and an analysis of the nature and frequency of all system activations.

These activities resulted in a list of success criteria, against which the success of the trial could be assessed, and were grouped under four broad categories of technical success, clinical efficacy, cost-benefit, and social/organizational. These criteria, described in detail by Hansen et al. (2013), differed markedly from those envisaged before the interactive process described above, and formed the basis of the final evaluation. While improved clinical outcomes are usually regarded as the primary benefit of telemedicine systems, in this case clinically driven activations of the system proved to be a relatively minor application, and the trial yielded too few such activations in any particular clinical category to achieve statistical significance. The way in which the success criteria were themselves outcomes of the combined process is shown in Figure 25.4, in which the vertical axis represents approximately a time axis.

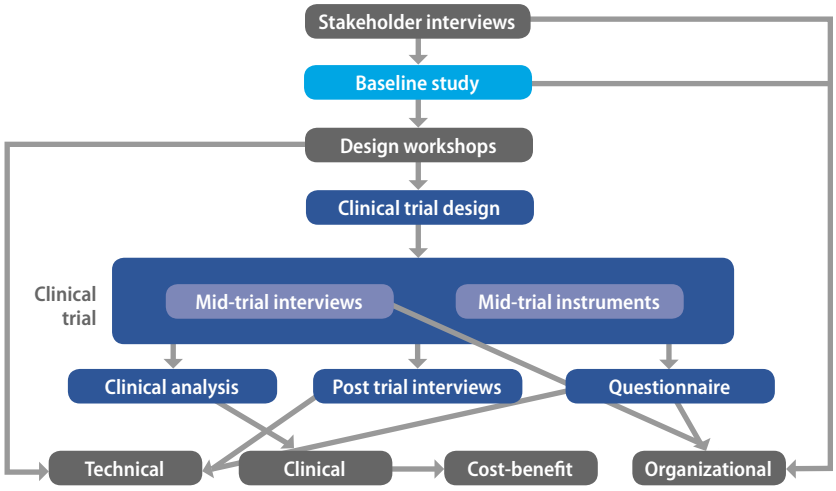


Figure 25.4. Components of the ECHONET project.

Note. From "Applying an integrated approach to the design, implementation and evaluation of a telemedicine system," by S. Hansen, L. Wilson, and T. Robertson, 2013, *Journal of the International Society for Telemedicine and eHealth, 1*(1), p. 27. Copyright 2013 by ISFTEH. CC BY License.

The success criteria and the measurable outcomes have been tabulated in Table 25.1. They are grouped as relating to the four broad categories of usability/technical, clinical, cost/benefit and organizational. Clinical benefits were difficult to quantify due to the diversity of clinical applications, but the validity of the technical solution was verified, and a range of social/organizational benefits were demonstrated, mainly among improved collegiate and educational interactions among the three participating sites.

It is clear from Table 25.1 that most of the perceived benefits were in the social/organizational area. However, the principal outcome of the project was a verification of the methodology of integrating design, implementation and eval-

uation processes. Many of the benefits were not envisaged at the beginning of the project, and the adaptive nature of the evaluation process ensured that these benefits could be assessed.

Table 25.1
Success Criteria for the ECHONET Project, Grouped Under Four Broad Evaluation Categories

Evaluation domain	Usability and technical	Clinical	Cost/benefit	Social/organizational
Success criteria (Evaluation criteria shaded)	Few faults	Reduced transfers		• Clinically safe; no adverse outcomes
	In routine use	More timely diagnosis	Continuing use following trial/clinical sustainable	• Number of bedside consults and number of participants
	Ease of use measured by number of users	Reduced travel for family and outpatients	Financially sustainable after trial	• Raising knowledge and skills (e.g., benchmarking ICU procedures at NWRH)
			Cost/benefit analysis based on other criteria outcomes	• Improved contact between ICUs (e.g., NWRH postings more popular) • Accepted as part of normal workflow (e.g., post-trial activations) • Strengthen ICUs (e.g., long-term benchmarking)

The most significant outcomes centred around improved collegiate relationships and educational opportunities among the users. Participants, in both the interviews and questionnaires, were very positive about the usability and usefulness of ECHONET, with some minor technical reservations. While all participants agreed that there were strong clinical benefits, the data sample was too small and diverse for this to be quantified by this study.

While the benefits of the collaboration supported by ECHONET for clinicians in the more remote hospital site at NWRH were more obvious and expected, clinicians in Hobart also recognized they had benefited from the collaborations made possible by the new technology. The educational benefits of ECHONET were realized early in the clinical trial. Education represents a good area in which to start using new telemedicine systems as sessions can be scheduled to allow familiarization with the system in a relatively low-pressure situation and routine use. The potential for ECHONET to be used for this purpose emerged early and strongly during the baseline study and this potential was confirmed and further explored during the clinical trial by clinicians at both hospitals.

25.5 Summary

This chapter has presented a view that Telehealth may be regarded as a "special case" in eHealth evaluation, in that it difficult to treat its components in isola-

tion from the context of usage. Nevertheless, typical telehealth evaluations tend to have focused on selected areas which include costs and resources, organizational and social aspects, and clinical benefits, rather than comprehensive coverage. Attempts to identify various sets of criteria, models and frameworks for evaluation have been described in the literature without achieving widespread consensus. These have been based around such disparate views as the inherent sequential characterization of a Telehealth intervention over time, or the taxonomic analysis of Telehealth along system functionality lines. It is argued that there is an overarching need to take a holistic approach and integrate different elements of evaluation to understand characteristics of the overall system of interest which is enabled by Telehealth. A case study has been presented to illustrate this process, borrowing from the central paradigm of participatory research as the holistic mechanism. This example was not intended to be definitive or exclude other approaches, but to emphasize the power of multifactor evaluations in such settings.

References

Agboola, S., Hale, T. M., Masters, C., Kvedar, J., & Jethwani, K. (2014). "Real-world" practical evaluation strategies: A review of telehealth evaluation. *JMIR Research Protocols, 3*(4), e75.

Bashshur, R. L. (1995). On the definition and evaluation of telemedicine. *Telemedicine Journal, 1*(1), 19–30.

Bashshur, R., Shannon, G., & Sapci, H. (2005). Telemedicine evaluation. *Telemedicine and e-Health, 11*(3), 296–316.

Bashshur, R., Shannon, G., Krupinski, E., & Grigsby, J. (2011). The taxonomy of telemedicine. *Telemedicine and e-Health, 17*(6), 484–494.

Bonell, C., Fletcher, A., Morton, M., Lorenc, T., & Moore, L. (2012). Realist randomised controlled trials: A new approach to evaluating complex public health interventions. *Social Science and Medicine, 75*(12), 2299–2306.

Brear, M. (2006). Evaluating telemedicine: lessons and challenges. *The Health Information Management Journal (Australia), 35*(2), 23–31. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/18209220

Chang, H. (2015). Evaluation framework for telemedicine using the logical framework approach and a fishbone diagram. *Healthcare Informatics Research, 21*(4), 230–238.

Doarn, C. R., Pruitt, S., Jacobs, J., Harris, Y., Bott, D. M., Riley, W., Lamer, C., & Oliver, A. L. (2014). Federal efforts to define and advance telehealth — A work in progress. *Telemedicine and e-Health, 20*(5), 409–418.

Clarke, M., & Thiyagarajan, C. A. (2008). A systematic review of technical evaluation in telemedicine systems. *Telemedicine and e-Health, 14*(2), 170–183.

Deshpande, A., Khoija, S., Lorca, J., McKibbon, A., Rizo, C., Husereau, D., & Jadad, A. J. (2009). Asynchronous telehealth: a scoping review of analytic studies. *Open Medicine, 3*(2), 69–91.

Ekeland, A. G., Bowes, A. S., & Flottorp, S. (2010). Effectiveness of telemedicine: A systematic review of reviews. *International Journal of Medical Informatics, 79*(11), 736–771.

Gagnon, M.-P., & Scott, R. E. (2005). Striving for evidence in e-health evaluation: lessons from health technology assessment. *Journal of Telemedicine and Telecare, 11*(suppl 2), 34–36.

Goran, S. F. (2012). Measuring tele-ICU impact: Does it optimize quality outcomes for the critically ill patient? *Journal of Nursing Management, 20*(3), 414–428.

Greenhalgh, T. (2012). Whole System Demonstrator trial: Policy, politics and publication ethics. *British Medical Journal, 345*, e5280.

Hailey, D., Jacobs, P., Simpson, J., & Doze, S. (1999). An assessment framework for telemedicine applications. *Journal of Telemedicine and Telecare, 5*(3), 162–170.

Hansen, S., Wilson, L., & Robertson, T. (2013). Applying an integrated approach to the design, implementation and evaluation of a telemedicine system. *Journal of the International Society for Telemedicine and eHealth, 1*(1), 19–29.

Hebert, M. (2001). Telehealth success: Evaluation framework development. *Studies in Health Technology and Informatics, 84*(2), 1145–1149.

Heinzelmann, P. J., Williams, C. M., Lugn, N. E., & Kvedar, J. C. (2005). Clinical outcomes associated with telemedicine/telehealth. *Telemedicine and eHealth, 11*(3), 329–347.

Hersch, W., Helfand, M., Wallace, J., Kraemer, D., Patterson, P., Shapiro, S., & Greenlick, M. (2002). A systematic review of the efficacy of telemedicine for making diagnostic and management decisions. *Journal of Telemedicine and Telecare, 8*(4), 197–209.

Hersch, W. R., Hickham, D. H., Severance, S. M., Dana, T. L., Pyle Krages, K., & Helfand, M. (2006). Diagnosis, access and outcomes: Update of a systematic review of telemedicine services. *Journal of Telemedicine and Telecare, 12*(suppl 2), 3–31.

Jackson, D. E., & McClean, S. I. (2012). Trends in telemedicine assessment indicate neglect of key criteria for predicting success. *Journal of Health Organization and Management, 26*(4), 508–523.

Jagosh, J., Macaulay, A. C., Pluye, P., Salsberg, J., Bush, P. L., Henderson, J., … Greenhalgh, T. (2012). Uncovering the benefits of participatory research: Implications of a realist review for health research and practice. *Milbank Quarterly, 90*(2), 311–346.

Jennett, P. A., Scott, R. E., Affleck Hall, L., Hailey, D., Ohinmaa, A., Anderson, C., … Lorenzetti, D. (2004). Policy implications associated with the socioeconomic and health system impact of telehealth: A case study from Canada. *Telemedicine and e-Health, 10*(1), 77–83.

Kidholm, K., Ekeland, A. G., Jensen, L. K., Rasmussen, J., Pedersen, C. D., Bowes, A., Flottorp, S. A., & Bech, M. (2012). A model for assessment of telemedicine applications: MAST. *International Journal of Technology Assessment in Health Care, 28*(1), 44–51.

Li, J., Wilson, L. S., Percival, T., Krumm-Heller, A., Stapleton, S., & Cregan, P. (2006). Development of a broadband telehealth system for critical care: Process and lessons learned. *Telemedicine and e-Health, 12*(5), 552–561.

Maeder, A., Gray, K., Borda, A., Poultney, N., & Basilakis, J. (2015). Achieving greater consistency in telehealth project evaluations to improve organizational learning. *Studies in Health Technology and Informatics, 209*, 84–94.

Marchal, B., Westhorp, G., Wong, G., Van Belle, S., Greenhalgh, T., Kegels, G., & Pawson, R. (2013). Realist RCTs of complex interventions – an oxymoron. *Social Science and Medicine, 94*(1), 124–128.

Nepal, S., Li, J., Jang-Jaccard, J., & Alem, L. (2014). A framework for telehealth program evaluation. *Telemedicine and e-Health, 20*(4), 393–404.

Scott, R. E., McCarthy, F. G., Jennett, P. A., Perverseff, T., Lorenzetti, D., Saeed, A., Rush, B., & Yeo, M. (2007). National telehealth outcome indicators project. *Journal of Telemedicine and Telecare, 13*(suppl 2), 1–38.

Stamm, B. H., & Perednia, D. A. (2000). Evaluating psychosocial aspects of telemedicine and telehealth systems. *Professional Psychology: Research and Practice, 31*(2), 184–189.

Steventon, A., Bardsley, M., Billings, J., Dixon, J., Doll, H., Hirani, S., … Newman, S., for the Whole System Demonstrator Evaluation Team. (2012). Effect of telehealth on use of secondary care and mortality: Findings from the Whole System Demonstrator cluster randomised trial. *British Medical Journal, 344*, e3874.

Sutcliffe, A., Thew, S., De Bruijn, O., Buchan, I., Jarvis, P., McNaught, J., & Proctor, R. (2010). User engagement by user-centred design in e-Health. *Philosophical Transactions of the Royal Society A. Mathematical, Physical and Engineering Sciences, 368*(1926), 4209–4224.

Tulu, B., Chatterjee, S., & Maheshwari, M. (2007). Telemedicine taxonomy: a classification tool. *Telemedicine and e-Health, 13*(3), 349–358.

Van Dyk, L. (2014). A review of telehealth service implementation frameworks. *International Journal of Environmental Research and Public Health, 11*(2), 1279–1298.

Wade, V. A., Kanon, J., Elshaug, A. G., & Hiller, J. E. (2010). A systematic review of economic analyses of telehealth services using real time video communication. *BMC Health Services Research, 10*, 233.

Wilson, L. S., Stevenson, D. R., & Cregan, P. (2009). Telehealth on advanced networks. *Telemedicine and e-Health, 16*(1), 69–79.

Wootton, R., & Hebert, M. A. (2001). What constitutes success in telehealth? *Journal of Telemedicine and Telecare, 7*(suppl 2), 3–7.

Part IV

Future Directions

Chapter 26

Building Capacity in eHealth Evaluation: The Pathway Ahead

Simon Hagens, Jennifer Zelmer, Francis Lau

26.1 Introduction

While much progress has been made in recent years, accommodating the growing demand for evidence relating to eHealth will require a continued focus on capacity building. Similar interest in evaluation capacity building extends into other disciplines and, as a consequence, has been the focus of discussion in the literature.

Labin, Duffy, Meyers, Wandersman, and Lesesne (2012) define evaluation capacity building as an "intentional process to increase individual motivation, knowledge and skills, and to enhance a group or organization's ability to conduct or use evaluation" (p. 308). For the purpose of this discussion, the focus will be on the broader health system's ability to conduct or use evaluation related to digital solutions. Preskill and Boyle (2008) describe the goal of evaluation capacity building as being "where members continuously ask questions that matter, collect, analyze, and interpret data, and use evaluation findings for decision-making and action" (p. 448). They go on to describe essential inputs including leadership support, incentives, resources and opportunities to transfer learning (Preskill & Boyle, 2008). This is consistent with the themes emerging from the capacity building experience in eHealth.

26.2 Motivation for Benefits Evaluation and Benefits Realization

Evaluation is a core component of an overall approach to benefits realization (Hagens, 2009). Clear and specific articulation of the benefits being targeted is

an important starting point. With this step, expectations can be set and the mobilization of required participants can begin. A next step is identification of key assumptions or conditions necessary for benefits to materialize, and action required to address them. These actions may be many and varied. Examples include decision support, user interface considerations, workflow or other process redesign, policy or practice change, or approaches to harvesting quality or productivity gains. A structured change management methodology can help ensure success. As part of this process, measurement against objectives allows for the opportunity to adapt and adjust based on the findings on an ongoing basis, thereby improving results. Information to manage course corrections and subsequent steps is always required. Stakeholders and funders will also want to know the value produced and have other accountability considerations addressed.

The most effective evaluations are managed with the end in mind, informed by the stakeholders who have the ability to apply the findings. Ideally, evaluation work will meet the needs of multiple stakeholders, and thus consider multiple perspectives. As discussed, funders and decision-makers are an important audience. Clinicians and other staff in clinical settings will also be interested. Evaluation can inform clinicians of progress achieved and help them get the most out of investments. Evidence to inform optimization of benefits is also critical for implementers and vendors. Academia supports knowledge translation through teaching and encourages rigour and quality of methods and analysis. The varied interests of stakeholder groups require consideration in the design and execution of evaluation. Meeting the needs of all stakeholders may require trade-offs. For example, formative or process evaluation can heavily inform adoption and optimization. Summative or outcome evaluation is required to effectively assess value.

As a result, many stakeholders need to have skills and assets to contribute to the evaluation process, but not all have an equal capacity to do so. Capacity to design, execute, and be responsive to evaluations is a top capacity need. Academia contributes substantially to addressing capacity needs and can be effective at publishing and communicating findings. Clinicians, health sector leaders, implementers, the vendor community, internal and external evaluators, and training providers can also play important roles. With growing needs for these skill sets, there is an opportunity for greater participation by all.

Effective evaluation also requires the focused engagement of those involved in digital health initiatives from users to implementation teams to leadership. For instance, time and support is needed to co-design evaluation frameworks, gain approvals, contribute insights, facilitate data collection, provide other input, and respond to evaluation findings.

26.3 The Foundation

Encouraging and supporting capacity development is best built upon a foundation of tried and tested frameworks, tools and processes. There are a number

of cross-sector structured approaches that have been applied to support benefits realization for digital health, such as Val IT (see http://www.isaca.org/knowledge-center/val-it-it-value-delivery-/pages/val-it1.aspx), Prosci (see https://www.prosci.com/) and value chains. There are also a number of tools that have been tailored to the health sector's needs. Some of important contributors to this body of knowledge and resources are discussed in greater depth in chapter 1 of this handbook.

In Canada, digital health-related resources have been developed by a variety of individuals and organizations. For instance, Canada Health Infoway's Benefit Evaluation Framework (discussed in detail in chapter 2) provides a high-level, coherent, evidence-based model to guide discussion of benefits and evaluation approaches (Lau, Hagens, & Muttitt, 2007). This framework, along with sets of indicators that focus on various types of digital health, have been regularly used to support measurement as part of a benefits realization cycle. The broader Clinical Adoption Framework is also a useful reference to consider the range of inputs influencing success (Lau, Price, & Keshavjee, 2011). Likewise, the Newfoundland and Labrador Centre for Health Information (NLCHI, n.d.) produced a range of materials including an evaluation framework and a series of successful evaluations as examples. Faculty at the University of Victoria's School of Health Information Science have also been productive, with a series of eHealth evaluation frameworks and tools through the jointly funded CIHR/Infoway eHealth Observatory (Lau, n.d.). In addition, a number of open and proprietary evaluation tools and frameworks are offered by solution vendors, consulting firms, and think tanks.

Internationally, there have also been many contributions. A notable health IT evaluation framework and toolkit was produced by the United States Agency for Healthcare Research and Quality (AHRQ, n.d.) to support their demonstration projects (Cusack & Poon, 2007). It was informed by some of the groundbreaking U.S. research which began emerging a number of decades ago. Another important contribution comes from the Organisation for Economic Cooperation and Development (OECD), which has been developing benchmark measures to allow comparison and knowledge sharing (OECD, 2013). They cover four major domains: provider-centric electronic records, patient-centric electronic records and services, health information exchange, and telehealth. While there are challenges with differing terminology and approaches to eHealth across countries, the OECD effort is proving important for supporting cross-national benchmarking and efforts by countries to enhance digital health measurement (Adler-Milstein, Ronchi, Cohen, Winn, & Jha, 2014).

Important foundational outputs of the work of organizations such as those discussed above also include practical tools to assist with conducting evaluations. The System & Use Assessment survey developed by Canada Health Infoway and its partners is one such example. It has been extensively applied across Canada over the last decade (Infoway, 2006, 2012). There are many other

similar examples of well-tested tools to make collection and interpretation of data easier for organizations building capacity.

Virtual communities have been important for sharing all of these resources, as well as the experiences of those involved. They also provide a forum to build fruitful relationships, such as connecting experienced evaluators with those in need of support.

While these resources provide a helpful starting point for those embarking on eHealth evaluation, there is an ongoing need for development and evolution. The rapid change of technology and its application in healthcare requires development of evaluation methodologies and tools to keep pace. Similarly, the growing demand for evidence to inform decision-making requires evaluation approaches aligned to evolving priorities and questions being posed. The increasing digitization of health has generated new data sources with substantial potential to improve evaluation options, as well as broadly inform the health system. Seizing this opportunity, however, takes careful planning and cooperation.

26.4 Approaches for Building Capacity

Just as the contributions to the knowledge base came from many different stakeholder groups, evaluation capacity building has come from across the sector, through leveraging evaluation expertise and capacity developed in other domains.

Basic undergraduate education through universities and colleges, a traditional approach for capacity building, has been impactful. While the University of Victoria offered the first health informatics program in Canada, there are now over 10 that train undergraduate students in the fundamentals of eHealth and its implications on the health system, and several that produce experts through their graduate programs. Fewer have courses specifically dedicated to evaluation.

More broadly, Canada's faculties of medicine, nursing, and pharmacy have undertaken specific initiatives to focus on how to better prepare students to practice in modern, technology-enabled, clinical environments (Baker, Charlebois, Lopatka, Moineau, & Zelmer, 2016). Supported by Infoway, the specific goals of this program were to:

- Ensure that clinicians-in-training are ready to practice in, and gain value from, an ICT-enabled environment when they graduate; and
- Integrate concepts and expectations related to the use of ICT in practice into curricula design and educational processes.

In a number of cases, these efforts include embedding competencies related to evaluation in health professional undergraduate education. Continuing education through academia and other education providers is also essential, as many professionals seek core skills to embark on evaluation work or to enrich their knowledge in key areas.

Recognizing the importance of capacity building and the critical role of academia, the Canadian Institutes of Health Research (CIHR) and Infoway partnered in 2008 to offer a five year CIHR-Infoway Chair in eHealth. This award, won by Francis Lau of the University of Victoria, proved a successful example of targeted funding making a significant impact, with outputs including some of the frameworks, publications, and communities referenced earlier (Lau, 2014).

Practical experience in undertaking and addressing the findings of evaluations is also important for building capacity, particularly given that the volume of evaluation activity has increased in recent years. This growth parallels the rise in evaluation in the public sector as a whole. Many investments in eHealth today have an explicit requirement for measurement, be it around the implementation process, the change effort, the adoption and/or the impacts. This was not previously the norm, but more sophisticated approaches to project delivery and an increasing demand for evidence-informed decisions has changed the expectations. With greater funding and attention from leadership, implementers have sought out evaluators from academia and the private sector, and often take the opportunity to grow in-house capabilities. Arguably, the most effective work comes from collaboration between these groups, matching those in a position to shape evaluations and generate knowledge with those who are in a position to apply the findings.

Growth in the volume of evaluation activity has required investments of financial, human, and other resources. Granting agencies have an important role in this area. CIHR has made some very important contributions over the past decade, with eHealth an explicit focus of a number of grant competitions and knowledge translation activities. Embedding evaluation as part of project plans and budgets is also increasingly common. Organizations delivering eHealth solutions are now more likely to require evaluation as a deliverable, and are able to budget for IT and engage skilled internal or external evaluators to support the work.

26.5 Approaches for Knowledge Translation and Benefits Realization

As important as increasing the capacity for conducting evaluations is increasing the application of findings. The Canada Health Infoway Benefits Evaluation Framework focuses on three purposes: accountability, informing clinicians and other digital health users, and driving benefits realization.

Accountability for investments made is increasingly important in the public sector and has been an important driver of expanded evaluation and performance management practices in Canada. Methodologies and reporting approaches must be tailored for this purpose. Clinicians, steeped in a culture of evidence-informed practice, similarly expect evidence to shape digital health design, implementation, and adoption, as well as its effective integration into clinical practice.

This includes supporting evidence-informed strategic planning and implementation. All stakeholders involved in implementation can benefit from evidence to inform optimization and realization of benefits. For instance, initial strategic planning typically includes a review of the evidence and critical success factors to inform priorities, assess options, and guide plans. Subsequently, evidence may help to drive enabling functionality like decision support, redesigning workflows to capture potential productivity improvements, addressing barriers to adoption like user interface challenges or inconsistent policies, or harnessing data for secondary use. While any of these factors may be identified during project planning, often the full value proposition emerges over time, with thoughtful observation, analysis, and ongoing response to feedback from users.

Traditional approaches to knowledge translation (KT), such as publications and conferences, remain central to the long-term objective of building a rich and robust knowledge base. They both enable communication to a range of audiences, and conferences increase the opportunity to build collaboration from that communication. Peer-reviewed literature helps to create quality standards that allow those applying the results to apply them appropriately and confidently. Limitations of peer review publication include delays (often in excess of a year), the effort required to complete the process, and disincentives for many outside the academic community to contribute findings.

In addition, KT approaches have been rapidly evolving, to both get evidence into the hands of decision-makers more quickly and encourage broad participation. Within specific projects, rapid cycle improvement methods can help to get actionable information into the hands of those with the ability to adapt plans and processes. Ideally, projects are designed with an optimization period. This ensures that resources are available to make adjustments as the process unfolds. Often quality improvement cycles are built into broader change management methodologies. The National Change Management Framework and supporting toolkit, developed by the Pan-Canadian Change Management Network with the support of Canada Health Infoway, positions evaluation as a central activity and provides some of the practical guidance required to enable long-term success (Infoway, 2013).

An expanding range of approaches beyond peer-reviewed journals are also being used to share knowledge across organizational boundaries. For instance, webinars, often tied to the kinds of communities described above, are increasingly prevalent and valuable. There are also well-regarded print/online journals and magazines, and growing online and social media options. Each of these has unique pros and cons, with considerations such as reducing disincentives to sharing experiences, streamlining process requirements and prerequisites, removing complexity to access information, and ensuring that the quality of information can be assessed by users. Integrated KT and multi-channel communications are important considerations.

26.6 Capacity Building Examples

This section provides selected examples of the capacity building outputs that are mentioned in the Foundation section (26.3) of this chapter. The examples cover peer support communities, knowledge and learning resources, and formal evaluation courses.

26.6.1 Peer Support Communities

Canada Health Infoway initiated a Benefits Evaluation community in 2007, as work was underway to put the evaluation strategy into operation. Early roadblocks had emerged in gaining buy-in from project teams to take accountability for evaluation and ensuring that there were people with the right skills to be successful. The community directly addressed these roadblocks, bringing stakeholder groups together and showcasing practical methodologies, effective partnerships, and the value of having evidence. Much credit for the early success of this community goes to the staff of the Newfoundland and Labrador Centre for Health Information, who brought substantial expertise to this forum and demonstrated the collaborative relationship they had achieved between implementers and evaluators (NLCHI, n.d.). Today, there is strong participation from many groups across Canada and the community contributed substantially to the development of a series of indicator sets, which are included in Infoway's Benefits Evaluation Technical report (Infoway, n.d.). It has evolved over the years to focus on emerging areas of need and to engage a broader audience. In addition, Canada Health Infoway frequently brings evaluation expertise into other Infoway-facilitated communities, like jurisdictional implementers groups, clinician reference groups or InfoCentral communities.

A further example is the virtual eHealth Benefits Evaluation Knowledge Translation (BE-KT) community, which evolved from the University of Victoria's (UVic) eHealth Observatory (Lau, n.d.). In 2012-13, researchers at the eHealth Observatory facilitated a virtual learning community in eHealth evaluation with a broad membership including implementers, policy-makers and academia. This community featured live online sessions with presentations from mentors, follow-up questions to prompt online discussions, and resources and links to support members in their evaluation activities (Bassi, Lau, Hagens, Leaver, & Price, 2013). The community attracted over 130 participants, many from outside academia, who were seeking the knowledge and network to increase the use of evaluation in their organizations. Over an 18-month period, the BE-KT community website was visited 4,425 times and viewed 14,683 times by both registered and unregistered members. Additionally during that period, 28 live seminar sessions were held on different topics related to eHealth evaluation. The presenters included researchers from the eHealth Observatory, Infoway benefits realization staff and jurisdictional representatives. The overall feedback from community members was largely positive, in that the effort had raised awareness of the importance of BE, where to find BE resources, and how to apply the

BE Frameworks, methods and tools. Interested readers can refer to the final report and lessons learned from the eHealth Observatory website (Bassi, 2014).

26.6.2 Knowledge and Learning Resources

Over the years, a growing number of online knowledge and learning resources on eHealth evaluation have been published. Examples of the organizations and groups that provide publicly available eHealth evaluation resources over the Internet are listed below.

- Canada Health Infoway maintains a rich repository of knowledge resources in benefits evaluation on its website (Infoway, n.d.). These resources include the Infoway BE Framework, the BE technical indicator report, and published jurisdictional BE reports in its online resource centre.

- The Newfoundland and Labrador Centre for Health Information has published the outputs of its benefits evaluation work done over the years on its website (NLCHI, n.d.). These resources include an inventory of published electronic health record (EHR) initiatives across Canada, a review of published EHR evaluation literature and reports, and a proposed evaluation framework for EHR initiatives. In particular the proposed framework describes a collaborative process working with stakeholders to develop meaningful and relevant evaluation study design and measures that can be implemented by healthcare organizations.

- University of Victoria eHealth Observatory: This is part of a five-year chair in eHealth award jointly funded by CIHR and Infoway to examine the effects of health information systems deployment in Canada. The website contains a set of eHealth evaluation frameworks, rapid evaluation methods and sample evaluation tools that can be applied and/or adapted in field evaluation studies of different eHealth systems (Lau, n.d.).

- The Agency for Healthcare Research and Quality was funded as part of the national strategy in the United States to improve the quality of care through IT. Over the years, the AHRQ Health IT website has amassed a rich set of resources that include health IT evaluation toolkits, AHRQ-funded health IT projects, published health IT evaluation studies and position papers in health IT adoption and evaluation (AHRQ, n.d.).

- Members of the European Federation of Medical Informatics (EFMI) working group on Evaluation (EVAL) and the International Medical

Informatics Association (IMIA) working group on Technology Assessment and Quality Improvement have published a set of guidelines for the reporting of evaluation studies in health informatics called STARE-HI (Talmon et al., 2009; Brender et al., 2013) and for good evaluation practice in health informatics called GEP-HI (Nykänen et al., 2011). These guidelines are invaluable resources that provide guidance on how one should design, conduct and report high-quality eHealth evaluation studies in the field setting.

- Organisation for Economic Cooperation and Development (OECD, 2013) offers model surveys and other benchmarking tools related to health information and communications technologies.

- Institute for Health Information Studies, UMIT — Researchers at the University for Health Sciences, Medical Informatics and Technology (UMIT) have published an online inventory of evaluation studies in medical informatics called the Web-based evaluation database or EvalDB (see Ammenwerth & de Keizer, 2005). This database contains over 1,800 published health IT evaluation studies and systematic reviews, and is updated on an ongoing basis. It is one of the most comprehensive inventories on eHealth evaluation studies published to date.

- The National Institutes of Health Informatics (NIHI) provides a suite of online education sessions, including a series on evaluation, with sections on qualitative and quantitative methods, that can be accessed at www.nihi.ca

26.6.3 Formal Evaluation Courses

The School of Health Information Science at the University of Victoria has been offering a graduate level course on eHealth evaluation since 2010 as part of its MSc program in health informatics. This course is delivered as a five-day intensive on-campus workshop with two weeks of online follow-up through Web-conference sessions. The course goals are to help students: (a) understand the types of evaluation frameworks, methods and studies available; (b) become knowledgeable in how evaluation studies are designed, conducted and reported; and (c) apply evaluation findings to inform healthcare policy and practice. The workshop is made up of class lectures and discussions, case studies, guest speakers, and individual and group assignments. The assignments provide students with opportunities to appraise published eHealth evaluation studies, and to apply best eHealth evaluation practice guidelines in eHealth case examples while designing an eHealth field evaluation study. The course covers (but is not limited to) the following topics:

- Methods of appraising and reporting eHealth evaluation studies (e.g., assessment of methodological quality, best practices in eHealth evaluation);
- eHealth evaluation frameworks (e.g., Infoway Benefits Evaluation Framework, Clinical Adoption Framework);
- eHealth evaluation study design and methods (e.g., quantitative versus qualitative, mixed methods, experimental, observational studies, surveys, usability studies); and
- examples of published eHealth evaluation studies (e.g., reviews, controlled and descriptive studies).

There are other Canadian universities that offer health-related evaluation courses as part of their graduate programs in eHealth. For example, students in the MSc eHealth program at McMaster University can enrol in such elective courses as Health Economics and Evaluation (C711), Fundamentals of Health Research & Evaluation Methods (HRM721), Economic Analysis for the Evaluation of Health Services (HRM737), and Approaches to the Evaluation of Health Services (HRM762). Students in the MSc of Health Informatics program at the University of Waterloo can enrol in the Evaluation of Public Health Program (PHS614) course as an elective. There is also an MSc program in Health Evaluation at the University of Waterloo with its entire curriculum focused on program evaluation in public health and health systems. Note that the courses mentioned at these universities are not necessarily specific to eHealth.

26.7 Looking Ahead

Some important opportunities emerge through exploring capacity building for evaluation. Partnerships between academia and such other stakeholders as implementation teams, clinical users, and funders, have proven so mutually beneficial as to warrant expansion. There is value in continuing to build, maintain, and share the pool of such resources as data collection tools and sample methodologies. Diversification of training opportunities from degrees to courses, workshops and online offerings, has been important for expanding the pool of evaluators. Integrating evaluation and optimization into the project life cycle has likewise proven valuable. Sharing and acting on the results of evaluation, both locally and more broadly, is also important, just as evidence-informed care has become the standard for clinical practice. Much progress has been made, but many opportunities remain to continue to build capacity in this domain.

References

Agency for Healthcare Research and Quality (AHRQ). (n.d.). *Health information technology*. Rockville, MD: Author. Retrieved January 20, 2016, from https://healthit.ahrq.gov/

Ammenwerth, E., & de Keizer, N. (2005). *Inventory of evaluation studies of information technology in health care: Trends in evaluation research, 1982 to 2002*. Tyrol, Austria: UMIT. Retrieved from https://evaldb.umit.at/

Baker, C., Charlebois, M., Lopatka, H., Moineau, G., & Zelmer, J. (2016). Influencing change: Preparing the next generation of clinicians to practice in the digital age. *Healthcare Quarterly, 18*(4), 5–7.

Bassi, J. (2014). *Increasing capacity in eHealth benefits evaluation: eHealth benefits knowledge translation community. Final report and lessons learned*. Toronto: Canada Health Infoway. Retrieved from http://ehealth.uvic.ca/community/2014.04.09-KT%20Community%20Full%20Report-v1.0.pdf

Bassi, J., Lau, F., Hagens, S., Leaver, C., & Price, M. (2013). Knowledge translation in eHealth: Building a virtual community. *Studies in Health Technology and Informatics, 183*, 257–262.

Brender, J., Talmon, J., de Keizer, N., Nykänen, P., Rigby, M., & Ammenwerth, E. (2013). STARE-HI: Statement on reporting of evaluation studies in health informatics, explanation and elaboration. *Applied Clinical Informatics, 4*(3), 331–358.

Cusack, C. M., & Poon, E. G. (2007). *Health information technology evaluation toolkit* (AHRQ Publication No. 08-0026-EF). Rockville, MD: Agency for Healthcare Research and Quality.

Hagens, S. (2009). Canadian EHR: Early benefits and the journey ahead. *Healthcare Information Management & Communications Canada*, *23*(4), 24–26.

Infoway. (2006). *Benefits evaluation survey process — System & use assessment survey*. Canada Health Infoway benefits evaluation indicators technical report. Version 1.0, September 2006. Toronto: Author. Retrieved from https://www.infoway-inforoute.ca/en/component/edocman/resources/toolkits/change-management/national-framework/monitoring-and-evaluation/resources-and-tools/991-benefits-evaluation-survey-process-system-use-assessment-survey

Infoway. (2012). *Benefits evaluation survey process — System & use assessment survey*. Canada Health Infoway benefits evaluation indicators technical report. Version 2.0, April 2012. Toronto: Author.

Infoway. (2013). *A framework and toolkit for managing eHealth change*. Toronto: Author. Retrieved from https://www.infoway-inforoute.ca/en/component/edocman/resources/toolkits/change-management/methodologies-and-approaches/1659-a-framework-and-toolkit-for-managing-ehealth-change-2

Infoway. (n.d.). *Benefits evaluation*. Toronto: Author. Retrieved from https://www.infoway-inforoute.ca/en/solutions/benefits-evaluation

Labin, S. N., Duffy, J., Meyers, D., Wandersman, A., & Lesesne, C. (2012). A research synthesis of the evaluation capacity building literature. *American Journal of Evaluation*, *33*(3), 307–338. doi: 10.1177/1098214011434608

Lau, F. (2014, May 5). Value of eHealth evaluation research in supporting healthcare reform in Canada [Web log post to *Infoway Connects*]. Ottawa: Canada Health Infoway. Retrieved from http://infowayconnects.infoway-inforoute.ca/2014/05/05/value-of-ehealth-evaluation-research-in-supporting-healthcare-reform-in-canada

Lau, F. (n.d.). *University of Victoria (UVic) eHealth observatory.* Victoria, BC: University of Victoria. Retrieved from http://ehealth.uvic.ca/index.php

Lau, F., Hagens, S., & Muttitt, S. (2007). A proposed benefits evaluation framework for health information systems in Canada. *Healthcare Quarterly, 10*(1), 112–118.

Lau, F., Price, M., & Keshavjee, K. (2011). From benefits evaluation to clinical adoption — Making sense of health information system success. *Healthcare Quarterly, 14*(1), 39–45.

Newfoundland and Labrador Centre for Health Information (NLCHI). (n.d.). *Health analytics: Benefits evaluation*. St. John's: Author. Retrieved from http://NLCHI.nl.ca/index.php/quality-information/health-analytics/benefits-evaluations

Nykänen, P., Brender, J., Talmon, J., de Keizer, N., Rigby, M., Beuscart-Zephir, M. C., & Ammenwerth, E. (2011). Guideline for good evaluation practice in health informatics (GEP-HI). *International Journal of Medical Informatics, 80*(12), 815–827.

Organisation for Economic Cooperation and Development (OECD). (2013). *OECD guide to measuring ICTs in the health sector*. Paris: Author.

Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation, 29*(4), 443–459. doi: 10.1177/1098214008324182

Talmon, J., Ammenwerth, E., Brender, J., de Keizer, N., Nykänen, P., & Rigby, M. (2009). STARE-HI: Statement on reporting of evaluation studies in health informatics. *International Journal of Medical Informatics, 78*(1), 1–9.

Chapter 27

Future of eHealth Evaluation

A Strategic View

Francis Lau

27.1 Introduction

In this handbook we have examined both the science and the practice of eHealth evaluation in different contexts. The first part of the handbook, on conceptual foundations, has provided examples of organizing schemes that can help make sense of eHealth as interdependent sociotechnical systems, and how these systems can be defined and measured. Depending on the purpose and scope of the planned evaluation, an eHealth system may be conceptualized under different assumptions and viewed from multiple lenses in terms of its makeup, behaviours and consequences. For example, an eHealth system may be evaluated in a narrow context at the micro level as an artefact for its technical performance under the information system quality dimension of the Benefits Evaluation Framework. Alternatively, the evaluation may take on a broader scope focusing on the macro-level governance, standards, and funding dimensions of the Clinical Adoption Framework.

The second part of the handbook concerns methodological details and has provided a collection of research approaches that can be applied to address different eHealth evaluation questions. They range from such quantitative methods as comparative and correlational studies to such qualitative methods as descriptive and survey studies. There are also methods that utilize both qualitative and quantitative data sources such as economic evaluation, modelling and data quality studies. In addition, there are published guidelines that can enhance the reporting quality of eHealth evaluation studies. The repertoire of such methods offers ample choice for the evaluator to plan, conduct, publish and appraise eHealth evaluation studies to ensure they are simultaneously rigorous,

pragmatic and relevant. The third part of the handbook, on selected eHealth evaluation studies, has provided detailed examples of field studies to demonstrate how the scientific principles of select eHealth evaluation frameworks and methods have been applied in practice within different settings.

The last part of the handbook on future directions addresses, first, the need to build capacity in eHealth evaluation and, second, the shifting landscape for eHealth evaluation within the larger healthcare delivery system. This final chapter of the handbook offers some observations on what this future may hold in the years ahead. This discussion is outlined under the topics of eHealth as a form of complex intervention, the need for guiding principles on eHealth evaluation methods, and taking a more strategic view of eHealth evaluation as part of the larger healthcare system. The chapter closes with some final remarks on key take-home messages on eHealth evaluation for readers.

27.2 eHealth as a Complex Intervention

There is growing recognition that healthcare interventions can be highly complex in nature. This can be due to the number of interacting components that exist in a given intervention, the types of behaviours required by those delivering and receiving the intervention, the number of targeted groups or organizations involved, variability in expected outcomes, and the degree of tailoring permitted in the intervention. Such complexity can lead to variable study findings and an apparent lack of tangible impact from the intervention (Craig et al., 2008).

According to Shcherbatykh, Holbrook, Thabane, and Dolovich (2008), eHealth systems are considered complex interventions since they are often made up of multiple technical and informational components influenced by different organizational, behavioural and logistical factors. The technical components include the eHealth system's hardware, software, interface, customizability, implementation and integration. The informational components include the operational logic, clinical expertise, clinical importance, evidence-based guidelines, communication processes and promotion of action. The organizational factors that can influence the system include its financing, management and training, the degree of vendor support, the stance of local opinion leaders, and feedback given and received. The behavioural factors include user satisfaction, attitudes, motivation, expectations, interdisciplinary interaction and self-education. The logistical factors include system design, workflow, compatibility, local user involvement, ownership, technological sophistication and convenience of access. Collectively these components and factors can interact in an unpredictable fashion over time to produce the types of emergent system functions, behaviours and consequences that are observed.

For complex eHealth interventions, Eisenstein, Lobach, Montgomery, Kawamoto, and Anstrom (2007) have emphasized the need to understand the intervention components and their interrelationships as prerequisites for effectiveness evaluation. These authors suggested that the overall complexity of an intervention can be a combination of the complexity of the problem being ad-

dressed, the intervention itself, inputs and outputs of the healthcare setting and the degree of user involvement. The group has developed the Oxford Implementation Index as a methodology that can be applied to eHealth evaluation (Montgomery, Underhill, Gardner, Operario, & Mayo-Wilson, 2013). This index has four implementation components that can affect intervention fidelity: intervention design; intervention delivery by providers; intervention uptake by patients; and contextual factors. These have been organized as a checklist to assess intervention study results. The checklist items are listed below.

- *Intervention design* – refers to core components of the intervention and the sequence of intended activities for the intervention group under study, as well as the usual practice activities for the control group.

- *Intervention delivery by providers* – refers to what is actually implemented which can be affected by staff qualifications, quality, use of system functions, adaptations and performance monitoring over time, such as the use of electronic preventive care reminders.

- *Intervention uptake by participants* – refers to the experience of those receiving the actual intervention that has been implemented, such as the patients who receive electronic preventive care reminders.

- *Contextual factors* – refers to characteristics of the setting in which the study occurs such as socio-economic characteristics, culture, geography, legal environment and service structures.

May and colleagues (2011) have proposed a Normalization Process Theory (NPT) to explain implementation processes for complex interventions in healthcare that can be extended to eHealth systems. The NPT has four theoretical constructs aimed to illuminate the embedding of a practice through what people actually do and how they actually work. These constructs are briefly described below (May et al., 2011, p. 2).

- *Coherence* – processes to understand, promote, or inhibit the intervention as a whole to its users. They require investments of meaning made by the participants.

- *Cognitive participation* – processes that promote or inhibit users' enrolment and legitimation of the intervention. They require investments of commitment by the participants.

- *Collective action* – processes that promote or inhibit the enactment of the intervention by its users. They require investments of effort made by the participants.

- *Reflexive monitoring* – processes that promote or inhibit the comprehension of the effects of the intervention. They require investments in appraisal made by the participants.

To translate NPT into practice, May et al. (2011) created an online survey as a Web-based toolkit to be completed by non-experts. The survey was field tested with 59 participants who responded to the questions and provided feedback to improve the content. The final version of the online survey has 16 statements where respondents can record their extent of agreement to each statement along a sliding bar from "completely agree" to "don't agree at all". See the Appendix for the 16 NPT statements and refer to the NPT website to access the toolkit (Normalization Process Theory [NPT], n.d.).

Mair and colleagues (2012) have conducted an explanatory systematic review to examine factors that promote or inhibit the implementation of eHealth systems using NPT as the organizing scheme. Of the 37 papers included in the review, they found there was little attention paid to: (a) work to make sense of the eHealth systems in terms of their purposes and values and to establish their value to users, and planning the implementation; (b) factors that would promote or inhibit stakeholder engagement and participation; (c) the effects on changing roles and responsibilities; (d) risk management; and (e) ways to reconfigure the implementation processes through user-produced knowledge. These findings suggest further work is needed to better understand the wider social framework and implications to be considered when introducing new technologies such as eHealth systems. The NPT may be a new and promising way to unpack the complexities associated with eHealth interventions that are currently not well addressed by traditional evaluation methods.

27.3 Guiding Principles for eHealth Evaluation Methods

There is a growing demand for governments and healthcare organizations to demonstrate the value of eHealth investments in ways that are rigorous and relevant. As such, eHealth evaluation is no longer considered an academic research activity but one that should be integral to the adoption of eHealth systems by healthcare organizations. As eHealth evaluation is increasingly being done by practitioners who may not be experienced in various evaluation approaches, there is an urgent need to ensure these evaluation studies are methodologically robust and reproducible. To explain and emphasize this need, Poon, Cusack, and McGowan (2009) have identified a set of common evaluation challenges faced by eHealth project teams funded by the Agency for Healthcare Research and Quality in the United States to deploy eHealth systems in their organizations.

These were mostly non-academic institutions with project teams that had a paucity of evaluation experience. The challenges found included having: evaluation as an afterthought; unrealistic evaluation scope and inadequate resources; a mismatch between the metrics chosen and the system being implemented; inadequate statistical power; limited data available; an improper comparison group; insufficient details on data collection and analysis; and an exclusive focus on quantitative methods.

There have been calls for the establishment of guiding principles to make eHealth evaluation more rigorous, relevant and pragmatic. For instance, Liu and Wyatt (2011) have argued for the need for more RCTs to properly assess the impact of eHealth systems. Rather than promoting the universal use of RCTs, however, they have pointed to the need for clarity on how to match study methods to evaluation questions. Specifically, an RCT is considered appropriate if there are significant costs and risks involved, since the study can answer questions on whether and how much an eHealth system improves practitioner performance and patient outcomes. Lilford, Foster, and Pringle (2009) have advocated the use of multiple methods to examine observations at the patient and system level, as well as the use of formative and summative evaluation approaches performed as needed by internal and external evaluators during different stages of the eHealth system life cycle. Similarly, Catwell and Sheikh (2009) have suggested the need for continuous evaluation of eHealth systems as they are being designed, developed and deployed in ways that should be guided by the business drivers, vision, goals, objectives, requirements, system designs and solutions.

Greenhalgh and Russell (2010) have offered an alternative set of guiding principles for the evaluation of eHealth systems. Their principles call for a fundamental paradigm shift in thinking beyond the questions of science, beyond the focus on variables, and beyond the notions of independence and objectivity. The argument being made is that eHealth evaluation should be viewed as a form of social practice framed and enacted by engaging participants in a social situation rather than a form of scientific testing for the sole purpose of generating evidence. As such, the evaluation should be focused on the enactments, perspectives, relationships, emotions and conflicts of participants that cannot be reduced to a set of dependent and/or independent variables to explain the situation under study. It also recognizes that evaluation is inherently subjective and value-laden, which is at odds with the traditional scientific paradigm of truth seeking that is purportedly independent and objective. In particular, these authors have compared these alternative paradigms under seven key quality principles described below (Greenhalgh & Russell, 2010, Table 1, p. 3).

- *Hermeneutic circle versus statistical inference* – Understanding of the situation through iterating between its different parts and the whole that they form rather than an adequately powered, statistical and representative sample from the population being studied.

- *Contextualization versus multiple interacting variables* – Recognizing the importance of context, its interpretive nature and how it emerges from a particular social and historical background rather than reliance on examining the relationships of a predefined set of input, output, mediating and moderating variables.

- *Interaction and immersion versus distance* – Focusing on engagement and dialogue between the evaluator and stakeholders and immersing in the socio-organizational context of the system under study rather than maintaining a clear separation for independence and objectivity.

- *Theoretical abstraction and generalization versus statistical abstraction and generation* – Relating observations and interpretations into a coherent and plausible model to achieve generalizability rather than demonstrating validity, reliability and reproducibility among study variables and findings.

- *Reflexivity versus elimination of bias* – Understanding how the evaluator's background, interests and perceptions can affect the questions posed, data collected and interpretations made rather than minimizing bias through rigorous methodological designs.

- *Multiple interpretations versus single reality amenable to scientific measurement* – Being open to multiple viewpoints and perspectives from different stakeholders rather than pursuing a single reality generated through robust study designs and methods.

- *Critical questioning versus empiricism* – There may be hidden political influences, domination and conflicts that should be questioned and challenged rather than assuming a direct relationship between the reality and the study findings based solely on the precision and accuracy of the measurements made.

From these quality principles we can expect different types of knowledge to be generated based on the underlying paradigms that guide the evaluation effort. For instance, under the traditional scientific paradigm we can expect the evaluation to: (a) employ objective methods to generate quantitative estimates of the relationships between predefined input and output variables; (b) determine the extent to which the system has achieved its original goals and its chain of reasoning; and (c) produce quantitative statistical generalization of the findings with explanatory and predictive knowledge as the end point.

By contrast, an evaluation under an interpretive/critical paradigm would tend to: (a) co-create learning through dialogue among stakeholders to understand

their expectations, values and framing of the system; (b) define the meaning of success through the struggles and compromises among stakeholder groups; and (c) provide a contextualized narrative with multiple perspectives on the system and its complexities and ambiguities (Greenhalgh & Russell, 2010, Table 2, p. 3).

27.4 A Strategic View of eHealth Evaluation

Since 2001 the Canadian federal government has invested $2.1 billion in eHealth through incremental and targeted funding allotments. Its provincial and territorial counterparts have also invested in cost-shared eHealth projects that included client and provider registries, interoperable EHRs, primary care EMRs, drug and lab information systems, diagnostic imaging systems, telehealth and consumer health. Despite such major investments, the evidence on eHealth benefits has been mixed to date (Lau, Price, & Bassi, 2014). Similarly, mixed findings are found in other countries as well. In the United Kingdom, progress toward an EHR for every patient has fallen far short of expectations, and the scope of the national programme for IT has been reduced significantly without any reduction in cost (National Audit Office [NAO], 2011). In the United States, estimated projected savings from health IT were $81 billion annually (Hillestead et al., 2005). Yet the overall results in the U.S. have been mixed. This may have been due to the sluggish adoption of eHealth systems that are neither interoperable nor easy to use, and the failure of healthcare organizations and providers to re-engineer their care processes, including provider payment schemes, in order to reap the full benefits of eHealth systems (Kellermann & Jones, 2013).

To guide eHealth policies, there is a need to expand the scope of eHealth evaluation beyond individual systems toward a more strategic view of where, how and in what ways eHealth fits into the broader healthcare system to demonstrate the overall return on value of the investments made. Kaplan and Shaw (2004) have suggested the evaluation of eHealth system success should extend beyond its technical functionality to include a mix of social, behavioural and organizational dimensions at a more strategic level that involve specific clinical contexts, cognitive factors, methods of development and dissemination, and how success is defined by different stakeholders. In order to evaluate these dimensions Kaplan and Shaw (2004, p. 215) have recommended 10 action items, which have been adapted as follows for this handbook:

1 Address the concerns of individuals/groups involved in or affected.

2 Conduct single and multisite studies with different scopes, types of settings and user groups.

3 Incorporate evaluation into all phases of an eHealth project.

4 Study failures, partial successes and changes in project definition or outcome.

5 Employ evaluation approaches that take into account the shifting nature of healthcare and project environment, including formative evaluations.

6 Incorporate people, social, organizational, cultural and ethical issues into the evaluation approaches.

7 Diversify evaluation approaches and continue to develop new approaches.

8 Conduct investigations at different levels of analysis.

9 Integrate findings from different eHealth systems, contextual settings, healthcare domains, studies in other disciplines, and work that is not published in traditional research outlets.

10 Develop and test theory to inform both further evaluation research and informatics practice.

In Canada, Zimlichman et al. (2012) have conducted semi-structured interviews with 29 key Canadian eHealth policy and opinion leaders on their domestic eHealth experiences and lessons learned for other countries to consider. The key findings are for eHealth leaders to emphasize the following: direct provider engagement; a clear business case for stakeholders; guidance on standards; access to resources for mid-course corrections of standards as needed; leveraging the implementation of digital imaging systems; and sponsoring large-scale evaluations to examine eHealth system impact in different contexts.

Similarly, at the 2011 American College of Medical Informatics (ACMI) Winder Symposium, a group of health informatics researchers and practitioners examined the contributions of eHealth to date by leading institutions, as well as possible paths for the nation to follow in using eHealth systems and demonstrating its value in healthcare reform (Payne et al., 2011). In terms of the role of eHealth in reducing costs and improving the quality of healthcare, the ACMI group suggested that eHealth systems can provide detailed information about healthcare, reduce costs in the care of individual patients, and support strategic changes in healthcare delivery.

To address the question of whether eHealth is worth the investment, the ACMI group have suggested the need to refocus the effort on more fundamental but strategic issues of what evidence is needed, what is meant by eHealth, what is meant by investment and how it is measured, and how we determine worth. These questions are briefly discussed below.

- *What evidence is needed?* Currently we do not routinely collect the data needed to help us determine the actual costs of eHealth systems and their economic and health impacts, including any unintended consequences. To do so on a continual basis would require structural changes to our healthcare operations and data models.

- *What is meant by eHealth?* We need to develop ways to articulate eHealth systems in terms of their functionality and co-factors that affect their design, deployment and use. Examples of co-factors include such areas as policies, process re-engineering, training, organization and resource restructuring, and change management. Also important is the recognition of the therapeutic dosage effect where there can be a differential impact with varying levels of eHealth system investment and adoption.

- *What is meant by investment and how it is measured?* We need to clarify who is making the investment, the form of that investment and the scope of the intended impacts. These can vary from the micro level that is focused on the burden and benefits for individual providers, to the macro level with a national scope in terms of societal acceptance of eHealth and its effects. For measurement, currently there are no clear metrics for characterizing the appropriate costs and benefits that should be measured, nor are there standardized methods for measuring them.

- *How do we determine worth?* While value is typically expressed in terms of dollars expended, productivity and effectiveness, we do not know what constitutes a realistic return on eHealth investments. This may depend on the initial states with respect to the level of investment made and the extent of eHealth system adopted. For example, with limited eHealth investment a healthcare organization may achieve only limited impact, whereas with a higher level of investment and broader stakeholder support one may achieve significant impact. For meaningful comparison these initial states may need to be normalized across studies and, given the small amount of evidence available to date, the focus should be on how to collect appropriate evidence in the future rather than pursuing a definitive answer on the worth of eHealth systems at this time.

27.5 Concluding Remarks

This chapter examined the future direction of eHealth evaluation in terms of its shifting landscape within the larger healthcare system, including the growing recognition of eHealth as a form of complex intervention, the need for alternate guiding principles on eHealth evaluation methods, and taking a more strategic view of eHealth evaluation as part of the larger system. This future should be built upon the cumulative knowledge acquired over many years in generating a better understanding of the role, makeup, behaviour and impact of eHealth systems through the application of rigorous methods in pragmatic evaluation studies that are relevant to multiple stakeholder groups. While there is still mixed evidence to date on the performance and impact of eHealth systems, the exemplary case studies provided throughout this handbook should offer some guidance on how leading healthcare organizations have planned, adopted and optimized their eHealth systems in order to reap tangible benefits over time.

In conclusion, the key messages for readers in terms of the future of eHealth evaluation and its implications within the larger healthcare system are summarized below.

- eHealth evaluation as an evolving science can advance our understanding and knowledge of eHealth as complex sociotechnical interventions within the larger healthcare system. At the same time, eHealth evaluation as a social practice can generate the empirical evidence needed to link the value of eHealth to the investments made from multiple stakeholder perspectives.

- There is a growing recognition of the need to apply theory-guided, multi-method driven and pragmatic design in eHealth evaluation that is based on best practice principles in order to build on the cumulative knowledge in health informatics.

- There is some evidence to suggest that, under the right conditions, the adoption of eHealth systems is correlated with clinical and health system benefits. Presently this evidence is stronger in care process improvement than in health outcomes, and the positive economic return is based on only a small set of studies. The question now is not whether eHealth can demonstrate benefits, but under what conditions can these benefits be realized and maximized.

References

Catwell, L., & Sheikh, A. (2009). Evaluating eHealth interventions: The need for continuous systemic evaluation. *Public Library of Science Medicine, 6*(8), e1000126.

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., & Petticrew, M. (2008). *Developing and evaluating complex interventions: new guidance*. Swindon, UK: Medical Research Council. Retrieved from http://www.mrc.ac.uk/documents/pdf/complex-interventions-guidance/

Eisenstein, E. L., Lobach, D. F., Montgomery, P., Kawamoto, K., & Anstrom, K. J. (2007). *Evaluating implementation fidelity in health information technology interventions*. In proceedings of the American Medical Informatics Association (AMIA) annual symposium, 2007, Chicago (pp. 211–215). Bethesda, MD: AMIA.

Greenhalgh, T., & Russell, J. (2010). Why do evaluations of eHealth programs fail? An alternative set of guiding principles. *Public Library of Science Medicine, 7*(11), e1000360. doi: 10.1371/journal.pmed.1000360

Hillestad, R., Bigelow, J., Bower, A., Girosi, F., Meili, R., Scoville, R., & Taylor, R. (2005). Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. *Health Affairs, 24*(5), 1103–1117.

Kaplan, B., & Shaw, N. T. (2004). Future directions in evaluation research: people, organizational and social issues. *Methods of Information in Medicine, 43*(3), 215–231.

Kellermann, A. L., & Jones, S. S. (2013). What it will take to achieve the as-yet-unfulfilled promises of health information technology. *Health Affairs, 32*(1), 63–68.

Lau, F., Price, M., & Bassi, J. (2014). Toward a coordinated electronic record strategy for Canada. In A. Cardon, J. Dixon, & K. R. Nossal (Eds.), *Toward a healthcare strategy for Canadians* (pp. 111–134). Montréal: McGill-Queen's University Press.

Lilford, R. J., Foster, J., & Pringle, M. (2009). Evaluating eHealth: How to make evaluation more methodologically robust. *Public Library of Science Medicine, 6*(11), e1000186.

Liu, J. L. Y., & Wyatt, J. C. (2011). The case for randomized controlled trials to assess the impact of clinical information systems. *Journal of the American Medical Informatics Association, 18*(2), 173–180.

Mair, F. S., May, C., O'Donnell, C., Finch, T., Sullivan, F., & Murray, E. (2012). Factors that promote or inhibit the implementation of e-health systems: an explanatory systematic review. *Bulletin of World Health Organization, 90*, 257–264. doi: 10.2471/BLT.11.099424

May, C. R., Finch, T., Ballini, L., MacFarlane, A., Mair, F., Murray, E., Treweek, S., & Rapley, T. (2011). Evaluating complex interventions and health technologies using normalization process theory: development of a simplified approach and web-enabled toolkit. *BMC Health Services Research, 11*, 245. doi: 10.1186/1472-6963-11-245

Montgomery, P., Underhill, K., Gardner, F., Operario, D., & Mayo-Wilson, E. (2013). The Oxford implementation index: a new tool for incorporating implementation data into systematic reviews and meta-analyses. *Journal of Clinical Epidemiology, 66*(8), 874–882.

National Audit Office. (2011). *The national programme for IT in the NHS: an update on the delivery of detailed care records systems*. London: Author. Retrieved from https://www.nao.org.uk/report/the-national-programme-for-it-in-the-nhs-an-update-on-the-delivery-of-detailed-care-records-systems/

Normalization Process Theory (NPT). (n.d.). *Implementing and evaluating complex interventions*. Retrieved from http://www.normalizationprocess.org/

Payne, T. H., Bates, D. W., Berner, E. S., Berstam, E. V., Covvey, H. D., Frisse, M. E., … Ozbolt, J. (2013). Healthcare information technology and economics. *Journal of American Medical Informatics Association, 20*(2), 212–217.

Poon, E. G., Cusack, C. M., & McGowan, J. J. (2009). Evaluating healthcare information technology outside of academia: observations from the National Resource Centre for healthcare information technology at the Agency for Healthcare Research and Quality. *Journal of American Medical Informatics Association, 16*(5), 631–636.

Shcherbatykh, I., Holbrook, A., Thabane, L., & Dolovich, L. (2008). Methodologic issues in health informatics trials: the complexities of complex interventions. *Journal of American Medical Informatics Association, 15*(5), 575–580.

Zimlichman, E., Rozenblum, R., Salzberg, C. A., Jang, Y., Tamblyn, M., Tamblyn, R., & Bates, D. W. (2012). Lessons from the Canadian national health information technology plan for the United States: opinions of key Canadian experts. *Journal of the American Medical Informatics Association, 19*(3), 453–459.

Appendix

NPT Survey Statements

(Source: May et al., 2011, pp. 8–9)

1 Participants distinguish the intervention from current ways of working

2 Participants collectively agree about the purpose of the intervention

3 Participants individually understand what the intervention requires of them

4 Participants construct potential value of the intervention for their work

5 Key individuals drive the intervention forward

6 Participants agree that the intervention should be part of their work

7 Participants buy into the intervention

8 Participants continue to support the intervention

9 Participants perform the tasks required by the intervention

10 Participants maintain their trust in each other's work and expertise through the intervention

11 The work of the intervention is allocated appropriately to participants

12 The intervention is adequately supported by its host organization

13 Participants access information about the effects of the intervention

14 Participants collectively assess the intervention as worthwhile

15 Participants individually assess the intervention as worthwhile

16 Participants modify their work in response to their appraisal of the intervention

Glossary

ACA: Affordable Care Act (United States)
ACE (inhibitor): angiotensin-converting-enzyme
ACMI: American College of Medical Informatics
ADE: adverse drug event
AHIMA: American Health Information Management Association
A1c Test: Glycated hemoglobin test
ANOVA: analysis of variance
AR: Action Research
ARI: acute respiratory infection
ART: antiretroviral therapy

BCMA: barcoded medication administration
BE: benefits evaluation
BE Framework: Benefits Evaluation Framework
BE-KT: Benefits Evaluation Knowledge Translation

CA Framework: Clinical Adoption Framework
CAD/DM: coronary artery disease and diabetes
CAMM: Clinical Adoption Meta-Model
CAS: complex adaptive system
CBA: cost-benefit analysis
CCA: cost-consequence analysis
CCS: clinical composite score
CD4: Cluster of Differentiation 4
CDM: chronic disease management
CDS: clinical decision support
CDSS: clinical decision support systems
CEA: cost-effectiveness analysis
CHEC: Consensus on Health Economic Criteria
CHEERS: Consolidated Health Economic Evaluation Reporting Standards
CHT: Canada Health Transfer
CI: clinical informatics
CIHR: Canadian Institutes of Health Research
CIS: Clinical Information Systems
CM: change management
CMA: cost-minimization analysis
CONSORT: Consolidated Standards of Reporting Trials

COPD: chronic obstructive pulmonary disease
CPCSSN: Canadian Primary Care Sentinel Surveillance Network
CPOE: computerized provider (or physician) order entry
CPR: cardiopulmonary resuscitation
CRC: colorectal cancer
CRCT: cluster randomized controlled trial
CSCW: Computer Supported Cooperative Work
CT: computerized tomography
CUA: cost-utility analysis
CUI: Common User Interface

DI: Diagnostic Imaging
DI/PACS: Diagnostic Imaging / picture archiving and communication system
DIS: Drug Information System
DPT: Data Presentation Tool
DQP: Data Quality Probe
DS: decision support
DVT: Deep Venous Thrombosis

ED: Emergency Department
EFMI: European Federation of Medical Informatics
EHR: Electronic Health Record
EMR: Electronic Medical Record
EMRAM: Electronic Medical Record Adoption Model
EQ-5D: EuroQol (European Quality of Life) five dimensions questionnaire
ER: Emergency Room

FMEA: Failure Modes and Effects Analysis
FSA: first specialist appointment
FTA: Fault Tree Analysis

GEP-HI: (guideline for) good evaluation practice in health informatics
GLMM: Generalized linear mixed model
GRADE: Grading of Recommendations Assessment, Development and Evaluation
GT: Grounded Theory
GUI: graphical user interface

HAPU: hospital-acquired pressure ulcer
HAZOP: Hazard and Operability
HbA1c: Hemoglobin A1c or glycated hemoglobin test
HCI: Human Computer Interaction
HFMEA: Healthcare Failure Modes and Effects Analysis
HIE: health information exchange

HIMSS: Healthcare Information and Management Systems Society
HIS: Health Information Systems
HIT: Health Information Technology
HL7: Health Level 7
HOIS: Hospital Information System
HSC: Health Sciences Centre (St. John's, NL, Canada)
HVDHB: Hutt Valley District Health Board (New Zealand)

ICBR: Incremental cost-benefit ratio
ICER: incremental cost-effectiveness ratio
ICT: Information and Communication Technology
ICU: intensive care unit
ICUR: Incremental cost-utility ratio
IMIA: International Medical Informatics Association
IS: Information Systems
ISO: International Organization for Standardization
ISPOR: International Society for Pharmacoeconomics and Outcomes Research
IT: Information Technology
ITIM: Information Technology Interaction Model
ITS: Interrupted time series

KP: Kaiser Permanente (United States)
KT: knowledge translation

MAR: medication administration record
MAST: Model for Assessment of Telemedicine Applications
MD: Medical Doctor
ME: medication error
MeSH: Medical Subject Headings
m-health: mobile health
MIS: management information systems
MMS: Multimedia Messaging Service
MRI: magnetic resonance imaging
MRSA: methicillin-resistant Staphylococcus aureus

NHS: National Health Service (United Kingdom)
NPT: Normalization Process Theory
NPV: net present value; also Negative Predictive Value (chapter 16)
NSCPP: National Shared Care Planning Programme (New Zealand)
NTI: narrow therapeutic index

ODEM: Ontario Diabetes Economic Model

PACS: picture archiving and communication system
PADE: possible adverse drug event
PAR: Participatory Action Research
PCHR: Personal Controlled Health Record
PCMH: Patient-Centered Medical Home
PCS: process composite score
PHI: personal health information
PHP: Personal Health Portal (Alberta, Canada)
PHR: personal health record
PHS: personal health services
PIMS: Patient Information Management System
PMS: practice management system
PP: Patient Protection
PPRNet: Practice Partner Research Network
PPV: Positive Predictive Value
PV: present value

QALY: quality-adjusted life year
QUERI: Quality Enhancement Research Initiative
QI: quality improvement
QOF: Quality and Outlooks Framework

RCT: Randomized controlled trial
RFP: request for proposals
RHIS: Routine Health Information Systems
RIS: Radiology Information System
RN: Registered Nurse
ROI: Return on investment
RREM: Rapid Response Evaluation Methods

SCMH: St. Clare's Mercy Hospital (St. John's, NL, Canada)
SCR: summary care record
SDLC: System development life cycle
SEM: Structural equation modeling
SMS: Short Message Service
SNOMED: Systematized Nomenclature of Medicine
STAMP: System Theoretic Accident Model and Processes
STARE-HI: statement on reporting of evaluation studies in health informatics
STROBE: Strengthening the Reporting of Observational Studies in Epidemiology
SUA: System and Use Assessment (survey tool)

TAM: Technology Acceptance Model
TAT: turnaround time

TOPSIS: Technique for Order Performance by Similarity to Ideal Solution
TRIP: Translation of Research into Practice

UKPDS: United Kingdom Prospective Diabetes Study
UMIT: the University for Health Sciences, Medical Informatics and Technology
UTAUT: Unified Theory of Acceptance and Use of Technology model

VDIS: Vermont Diabetes Information System

WSD: Whole Systems Demonstrator

XML: Extensible Markup Language

About the Contributors

- **Reza Alaghehbandan**, MD, MSc, is a Surgical Pathologist and a Clinical Epidemiologist with the Faculty of Medicine at the University of British Columbia. His primary research interest is focused on oncologic surgical pathology spanning from use of novel biomarkers to the development of external quality assurance and proficiency testing in breast cancers. Reza's other longstanding research interests include injury and chronic disease epidemiology as well as health issues related to vulnerable populations (i.e., women/children and Indigenous peoples) from human rights and social justice perspectives. Reza also has extensive expertise in developing and utilizing administrative health databases and EHR/EMR data for conducting population health research and surveillance.

- **James G. Anderson**, PhD, earned a BES in Chemical Engineering, MSE in Operations Research and Industrial Engineering, MAT in Chemistry and Mathematics, and a PhD in Education and Sociology from the Johns Hopkins University. He is the former Director of the Division of Engineering of the Evening College at Johns Hopkins University. At Purdue University, he has served as Assistant Dean for Analytical Studies of the School of Humanities, Social Sciences and Education (1975 to 1978), Associate Director of the Health Services Research Training Program supported by the U.S. Public Health Service (1971 to 1976), Director of the Social Research Institute (1995 to 1998), and Co-Director of the Rural Center for AIDS/STD Prevention (1994 to the present day). He is the author or co-author of five books.

- **Paule Bellwood**, MSc, is a PhD candidate enrolled in the Experimental Medicine program at University of British Columbia. Her research is focused on health informatics, targeting consumer health informatics in particular. Paule's research interests include end user involvement in and processes of health information and consumer health systems design.

- **Simon Diemert**, BSEng, is a graduate student studying Computer Science at the University of Victoria. His research interests include software engineering in healthcare (specifically focused on medications management) and engineering for safety critical systems.

- **Jeff Dowden**, BSc, MAS, is a statistician with the Newfoundland and Labrador Centre for Health Information. He has over 15 years combined educational and work experience in the areas of statistical analysis/research of health data; data analytics; secondary uses of administrative data; analysis of large data; and project leadership to support policy development. Jeff has co-authored a number of articles in the areas of statistical methods in health data and population health and has presented at national conferences. Jeff has been a methodological lead/co-lead in various projects to help enhance population health in the Province of Newfoundland and Labrador.

- **Rong Fu**, PhD, is an Assistant Professor in Sociology at Siena College. She received her dual-title PhD in Sociology and Gerontology from Purdue University, West Lafayette, Indiana. Her fields of study include medical sociology, health, and social gerontology. She has extensive training in quantitative methods and has teaching experience in social statistics and research methods. She has engaged in independent and collaborative research projects related to social behaviours, health inequality, and aging, among others. Her dissertation examined how childhood and later life stressors affected cognitive health in older adults.

- **Veeresh Gadag** has a PhD in the field of Applied and Biostatistics. His field of expertise is in the field of application of biostatistics to the health-services-research, continuity of care, adverse drug events, epidemiology of chronic diseases such as cancer, diabetes, cardiovascular diseases. He has been also working in the area of population health as it pertains to seniors' age friendly communities. In the area of theoretical statistics, he has done research in stochastics processes, time series, risk analysis, branching processes, the detection of outliers, and sampling designs, among others.

- **Yulong Gu**, PhD, is an Assistant Professor of Health Science at Stockton University, New Jersey. Before joining Stockton University, she worked as a postdoctoral researcher at the University of Auckland from 2011 to 2015. She has worked on a range of research projects linked to grants and New Zealand National Health Information Technology Board contracts, which investigated subjects including electronic referral systems, shared care, personal health records, mobile health and medication adherence. She has a strong interest in knowledge management, health IT evaluation, grounded theory and thematic analysis.

- **Ryan Habibi**, HBSc, is a graduate student studying Computer Science at the University of Victoria. Ryan's research focuses on human computer interaction and usability for data privacy in the healthcare domain.

- **Simon Hagens**, MBA, leads evaluation at Canada Health Infoway. In his 10 years with Infoway, Simon's largest contributions have been the development and implementation of the Benefits Evaluation strategy and the establishment adoption requirements to drive adoption. Prior to Infoway, Simon has been a manager in a primary care and community health organization, where he led the implementation of an electronic medical record system. As a market researcher in the pharmaceutical industry, Simon applied a variety of qualitative and quantitative research approaches to provide intelligence and recommendations for marketing major products. Simon holds a BSc from the University of Guelph and an MBA from McGill University, and has over 15 years of healthcare experience.

- **Anne Holbrook**, MD, PharmD, MSc, FRCPC, is Director of the Division of Clinical Pharmacology & Toxicology and a Professor in the Department of Medicine at McMaster University. She is also an Active Medical Staff member at St Joseph's Healthcare, Hamilton and Hamilton Health Sciences. Her clinical practice combines Clinical Pharmacology and Internal Medicine and specializes in the care of patients with multiple diagnoses and complex medication regimens. She has been a drug policy advisor to federal, provincial and regional governments for decades, influencing budgets of more than $5 billion per year. Dr. Holbrook is currently an investigator on multiple peer-reviewed, publicly funded research teams or projects, including a program of eHealth research — all oriented to improving the quality and cost-effectiveness of prescribing and medication use.

- **Spyros Kitsiou**, PhD, is an Assistant Professor of Biomedical and Health Information Sciences at the University of Illinois at Chicago and an Editorial Board member of the *Journal of Medical Internet Research*, mHealth and uHealth. His current research interests focus on the effectiveness of home telemonitoring and mHealth interventions for patients with chronic conditions as well as the conduct of systematic reviews, meta-analyses, and overviews of systematic reviews to support evidence-based practice in health informatics. His publications have appeared in the *Journal of Medical Internet Research*, *Information and Management, International Journal of Healthcare Technology and Management,* and at several international conferences.

- **Craig Kuziemsky**, PhD, is an Associate Professor and University Research Chair in Healthcare Innovation at the Telfer School of Management, University of Ottawa. In 2013 he was named The University of Ottawa Young Researcher of the Year in the Humanities and Social Sciences. Dr. Kuziemsky's research focuses on developing new approaches for modeling collaborative healthcare delivery so we can better design information and communication technology (ICT) to support different contexts of collaboration. His work has defined the structural aspects necessary to support collaboration as well as the behavioural and social processes that shape how the structural components work. His studies of collaboration have used concepts such as complexity theory to understand the nature of collaborative interactions in different healthcare settings (clinical healthcare and public health for disaster management).

- **Francis Lau**, PhD, is a Professor in the School of Health Information Science at the University of Victoria Canada. He has a PhD in medical sciences with specialization in medical informatics. He has a diverse background in business, computer and medical sciences, with 14 years of professional experience in the health IT industry. Dr. Lau's research foci are in health information system evaluation, clinical vocabularies, and palliative/primary care informatics. During 2008-2013 he was the recipient of the eHealth Chair funded by Canadian Institutes for Health Research / Canada Health Infoway to establish an eHealth Observatory to examine the impact of health information system deployment in Canada.

- **Don MacDonald**, PhD, is the Vice-President of the Analytics and Evaluation Services Department with the Newfoundland and Labrador Centre for Health Information. In addition to his role in numerous applied health research studies, Don was the lead investigator for the several studies evaluating the benefits of the province's EHR. In 2007, Don was named the benefits evaluation Subject Matter Expert for Picture Archiving and Communications Systems by Canada Health Infoway. He was chosen to develop Infoway's evaluation indicators for the diagnostic imaging, and helped create Infoway's EHR National Benefits Evaluation Framework. A graduate of Memorial University of Newfoundland and Labrador, Don obtained his PhD in the area of EHR benefits evaluation. Don also has an undergraduate degree in Mathematics and Statistics, a Diploma in Clinical Epidemiology, a Master's in Science (Medicine). His current work focuses on developing a provincial health analytics model in support of health system management.

- **Anthony J. Maeder**, PhD, joined Flinders University in April 2016 as Chair in Digital Health Systems from Western Sydney University, where he founded the Telehealth Research and Innovation Laboratory in 2010. He was previously appointed as Research Director of the CSIRO Australian eHealth Research Centre in Brisbane, from 2004 to 2008. He is a Fellow of the Institution of Engineers Australia and a Fellow of the Australian Computer Society. He is currently a member of the Standards Australia IT-14 Health Informatics Committee. Professor Maeder was president of the Australasian Telehealth Society in the period 2010-2012, and joined the Governing Board of the International Society for Telemedicine and eHealth in 2013. In 2015 he was appointed as the Fulbright Distinguished Professor in Life Sciences and Agriculture at Kansas State University.

- **Lynne Nemeth**, PhD, is a Professor in the College of Nursing at the Medical University of South Carolina, a Nurse Researcher at the Medical University Hospital Authority and a PPRNet Investigator. She has extensive clinical, educational, administrative and research experience. Her current and recent research is focused on translation of research into practice. She is an expert qualitative researcher with an emphasis on process evaluation and implementation research. She currently collaborates on research implementing mobile health applications in chronic care populations, and community engaged research. She teaches Advanced Qualitative and Mixed Methods Research to PhD students and chairs or mentors numerous PhD dissertations.

- **Doreen Neville** has a BN (Nursing, MUN, 1979), a MSc (Nursing, Toronto, 1984), a MSc (Health Services Administration, Harvard, 1986), and a ScD (Health Policy and Management, Harvard) and holds a faculty appointment as Associate Professor, Health Policy and Health Care Delivery, Division of Community Health and Humanities, Faculty of Medicine, Memorial University. Her research interests lie in the area of applied health policy, namely in the evaluation of the effectiveness of health information systems and health system governance models. Her teaching interests include the use of technology to enhance achievement of learning outcomes and the creation of teaching and learning environments to maximize engagement among educators, students and key stakeholders in the community. In October 2011, Dr. Neville was appointed to a second term as Associate Vice-President (Academic), with specific responsibility for Planning, Priorities and Programs.

- **Guy Paré**, PhD, is a Professor of Information Technology Management and holds the Chair in Information Technology in Health Care at *HEC* Montréal. His current research interests involve the clinical, organizational and economic impacts of telehealth applications, electronic medical record systems and health information exchange projects as well as the barriers to adoption and implementation of health IT. His work has appeared in top-ranked scientific journals and presented at several international and national conferences. Professor Paré's expertise in e-health has been sought by numerous organizations including the World Health Organization, the Department of Health in France, Canada Health Infoway, and the Québec Ministry of Health and Social Services. He was elected a member of the Royal Society of Canada in 2012.

- **Morgan Price**, MD, PhD, CCFP, FCFP is an assistant professor (University of British Columbia, Department of Family Practice, Island Medical Program) and family physician practicing in the inner city with underserved populations in Victoria. Dr. Price is an Adjunct Professor in Computer Science and Health Information Science at the University of Victoria. His area of interest is in design and adoption research of clinical information systems, focused in primary care, decision support, and consumer oriented information systems (Personal Health Records). He has been Primary Investigator on several successful clinical informatics design and research projects exploring clinical decision support, EMR adoption, personal health records, and medication information systems.

- **Khokan C. Sikdar**, PhD, is a Lead Biostatistician in the Libin Cardiovascular Institute of Alberta, University of Calgary. His works centres on developing and applying innovative biostatistical and epidemiologic methods to advance clinical and public health sciences. Dr. Sikdar provides leadership and recommendation in statistical consultation and analytics for members of the institute and oversees the operations of the institute's Biostatistics Analysis Core Lab. He leads the surveillance team of Alberta's Strategy for Patient Oriented Research, is an adjunct professor in the Community Health Sciences at University of Calgary and a full member of O'Brien Institute for Public Health. Over the years of his professional career, he has been involved in more than 100 projects in a variety of health research areas, including studies that focus on cardiovascular disease, colorectal cancer, diabetes, hypertension, and adverse drug reactions.

- **Peter Wang**, PhD, is a tenured professor (Epidemiology) in the Division of Community Health and Humanities, Faculty of Medicine, Memorial University of Newfoundland (MUN). Dr. Wang graduated from Tianjin Medical University, China and received his PhD in Epidemiology from the Dalla Lana School of Public Health, University of Toronto. Dr. Wang has broad research interests including chronic disease epidemiology, lifestyles and health, immigrants' health, statistical modelling and genetic epidemiology. Dr. Wang has been a lead investigator or a co-investigator for more than 30 research grants that have been supported by various provincial, national and international funding agencies. Dr. Wang is a reviewer for a number of funding agencies and academic journals; he has published more than 80 articles in peer-reviewed scientific journals.

- **Jim Warren**, PhD, is Professor of Health Informatics at the University of Auckland. After receiving his BS and PhD from University of Maryland, he was with the University of South Australia from 1993 to 2005. From 2008 to 2010 he served a term as Chair of Health Informatics New Zealand. He is also a founding Fellow of the Australasian College of Health Informatics. Jim's primary research interest is in IT for chronic condition management, whether this is through improved "business intelligence" (or data mining), clinical decision support tools for health providers, or information systems to better empower health consumers in their own care.

- **Jens H. Weber**, PhD, PEng, is a Professor of Software Engineering in the Faculty of Engineering, at the University of Victoria, British Columbia, Canada. He has an Adjunct Faculty appointment in the University's School of Health Information Science and is an Adjunct Professor at the University of British Columbia, Faculty of Medicine, Department of Family Practice. He is licensed as a practicing Professional Engineer in the Province of British Columbia.

- **Laurence S. Wilson** received a PhD in Cosmic Ray Physics in 1973 and after brief periods in research and teaching roles at the Universities of Oxford and Sydney respectively, established an international reputation for research in medical imaging in the Ultrasonics Laboratory of Australia's Commonwealth Scientific and Industrial Research Organisation (CSIRO). Particular areas of interest includes ultrasonic tissue characterisation, Doppler, and knowledge-based imaging. He later led teams in home telecare and broadband telehealth for critical care, as well as leading CSIRO's health-based ICT program. Other research interests include the role of Human Factors in telehealth. He has retired from full time work and is currently an Adjunct Professor at Western Sydney University.

- **Jennifer Zelmer**, PhD, is President of Azimuth Health Group, Editor-in-Chief of *Healthcare Policy*, an Adjunct Faculty Member at the University of Victoria, and serves on a range of health-related advisory committees and boards. Previously, Dr. Zelmer held a series of progressively responsible leadership positions in Canada and abroad, mostly recently as Executive Vice President at Canada Health Infoway, as Chief Executive Officer of the International Health Terminology Standards Development Organization, and as Vice President of Research and Analysis at the Canadian Institute for Health Information. She received her PhD and MA in economics from McMaster University and her BSc in health information science from the University of Victoria.

Victoria, BC Canada
press@uvic.ca

ISBN 978-1-55058-601-5