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:

- **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 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-

---

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 shortage 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


