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 subpro-
cesses (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 evalua-
tion 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 “log-
ical 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 ob-
server. 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 con-
trolled trial (RCT) research design that predominated in medicine for the eval-
uation of drugs and therapies. HIT was thought to be another intervention that
could be measured and evaluated using controlled conditions to isolate a par-
ticular intervention. However, over time it was shown that these approaches do
not work well for evaluating the complex multifaceted nature of eHealth im-
plementation (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 conse-
quences (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 be-
\textit{\textless\textless}ond order entry itself, such as communication and workflow issues, changes
in the power structure, and the creation of new work. These unintended con-
sequences emphasized the fact that the evaluation of HIT must go beyond just
the objective process being automated to also consider the contextual environ-
ment 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 method-
ologies and approaches including qualitative, quantitative and mixed methods
approaches. Quantitative approaches are useful when we want to evaluate spe-
cific aspects of an information system that are independent, objective, and dis-
crete 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 con-
ducted 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 ob-
jective 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 con-
text of the system being studied, and the specific events and processes that de-
fine 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 under-
standing 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 sys-
tem and allow the construction of models based on the evaluation. These mod-
els then serve as theories that can be tested using quantitative approaches. An
example of mixed methods research in eHealth evaluation is the aforemen-
tioned 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 quan-
titative approaches both to determine frequencies of the unintended conse-
quences and to compare frequencies across different settings.

While mixed method approaches can be a useful approach for eHealth eval-
uation 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, Mamlín, 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 healthcare (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
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 non-linear 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


