

24

Evaluation of Health Information Systems

Hamish S. F. Fraser

Take-Home Messages

- Health information systems (HISs) are increasingly used in low- and middle-income countries, but there is still limited data on their performance and impact.
- Evaluation studies need to examine the whole process of creation of an HIS, from requirements gathering and workflow analysis through development, initial deployment, scale-up and longer-term use, and focus on the needs of a range of stakeholders.
- Individual evaluation studies usually focus on specific aspects of HISs, but may be linked into a multistage evaluation, for example assessing performance, impact, and costs.
- It is important to be aware of the many potential biases in evaluation of HISs, and how these can be minimized with appropriate study designs.
- Rather than focus on a specific technology, studies should be designed to evaluate the underlying principles of HIS interventions, such as access to key data at point of care.
- Data quality is a key issue for all HISs and its regular monitoring and evaluation should be included in all projects.
- While large-scale randomized controlled trials may be required to assess clinical impact, smaller, low-cost studies can provide critical data for design, implementation, and scale-up of systems and help in the design of larger studies.
- Initial results from well-designed clinical trials show that HIS use in low- and middle-income countries can improve key aspects of care, including enrollment and retention in care; access to key laboratory data; time to initiation of lifesaving treatment, adherence to treatment, drug requirement forecasting, and effective disease surveillance.

Summary

Health information systems (HISs) have great potential to improve access to care, quality of care, and management of health systems, as well as impact general health and healthy lifestyles. Despite this potential, there is to date limited evidence that HISs do improve care, especially in low- and middle-income countries (LMIC). Evaluation of HISs builds on the techniques and study designs typically used for other health interventions such as medications, but there are key differences. HISs are usually embedded in the health system and the workflow of patient care, and therefore represent complex interventions that operate through interaction of technology, people, organizations, and infrastructure. Despite these challenges, effective evaluation is possible, and an increasing number of well-designed studies of HISs in LMIC have been published in the last decade. In this chapter, we describe the basic principles of evaluation of HISs and provide examples of effective studies. We review the distinction between formative and summative studies and between quantitative and qualitative studies. We will also briefly discuss qualitative research methods and economic evaluation and discuss examples of effective studies in LMIC. Finally, we will discuss key issues, including data-quality measurement and improvement and the Learning Health System. We also provide a short list of further reading for those interested in learning further about carrying out studies. Readers are encouraged to read chapter 5, “Modern Epidemiology and Global Health in the era of Information Systems and mHealth.”

Introduction

Why Evaluate Health Information Systems?

HISs are becoming increasingly widely used worldwide, including in some of the lowest-income countries such as Haiti, Rwanda, and Cambodia. As discussed in earlier chapters, there are many good reasons why these systems are growing in terms of perceived need, improvements in software, and in the availability of smaller, cheaper, more capable hardware that uses a lot less power. However, the deployment and use of HISs should be based on real information needs and on solid evidence that their use improves aspects of health care and patient health more generally. This requires rigorous evaluation studies and publication of evidence to allow development of robust guidance for decision makers. Evaluation studies need to assess a range of parameters—including system performance, usability, stability, and, crucially, data quality—as well as the impact of the systems on patient-care processes and, ideally, clinical outcomes; and finally also examine the costs and cost-effectiveness of such systems. In this chapter, we will discuss the types of evaluations that can be performed and the different evaluation methods, with their strengths and weaknesses, as well as give examples of some key evaluation results.

An important point to note is that most HISs do not undergo evaluation at all stages of development, and frequently are not evaluated at all. If the technology used is mature, there may not be a requirement to carry out the evaluation of the earlier stages of software development. However, if there are significant changes, it is important to ensure they have been implemented correctly and assess their impact. Evaluations do not need to be large or complex to yield valuable insights and data that can help the development and implementation process and inform decision makers [1]. A goal of this chapter is to illustrate that evaluation can be carried out by people with a wide range of backgrounds and in a range of environments without necessarily requiring large resources.

Types of Evaluation

HISs are often described as sociotechnical systems—their effect is based on the often-complex interactions of technology, people, organizations, and infrastructure. Therefore, evaluation of such systems must examine the technology, the human elements, and the interactions between them. In practice, this is reflected in study methods, including the division into *quantitative methods* and *qualitative methods*, with the former focusing on numeric measurement of effects, and the latter particularly focusing on the experiences and viewpoints of users.

It is important to address the needs of varied users and to examine systems at each stage, from original requirements gathering to long-term performance. Box 24.1 shows a summary of these stages. In *stage one*, problem definition includes requirements gathering and translation into design and architecture. Mistakes here will clearly impact the ability of the system to address user needs. In *stage two*, bench testing addresses the issue of ensuring the software and hardware combination is stable, usable, and ideally is shown to work well with authentic test data. *Stages three to five* are more conventionally thought of as HIS-evaluation targets, but the distinctions are important here. *Stage three* is based on observing the use of systems in real environments. This is particularly important for environments with poor infrastructure, as often seen in LMIC. Systems need to be *functional* and *used* regularly by staff, and key functions like reports and decision-support tools need to function effectively. *Stage four* is where typical impact studies take place, such as randomized controlled trials. *Stage five* represents monitoring the longer-term use of the system once deployed. With the rapid changes in health systems and technology, a system that works in the initial implementation will not necessarily continue to do so longer term. In addition, the roll-out of a system to new sites often throws up new problems that need to be detected and addressed.

This had led to the development of a variety of evaluation frameworks, one of which is shown in box 24.1.

Box 24.1**Five Levels of Evaluation [2]**

1. Problem definition
2. Bench testing
3. Field trials: Observational
4. Field trials: Interventional
5. Long-term follow-up

Formative and Summative Evaluation

A critical first stage in implementing and optimizing an HIS is ensuring that it is developed to address a real need and ensuring that its design and user interface are appropriate. (These issues are discussed in more detail in chapter 4 and chapter 21.) The next stage is ensuring that it functions as expected. This includes stability, speed and responsiveness, usability, usage, and data quality, tested in the real environment with real users, and represents stage three or four in box 24.1. This type of study is often referred to as a *formative evaluation*. It is typically carried out by, or in partnership with, the system developers and results in improvement to the system or implementation. These types of evaluations typically include quantitative and qualitative methods (see below for explanation). An example of a formative evaluation of electronic health record (EHR) system performance is given in the “Data Quality Improvement” section, below.

In contrast, an evaluation study intended to assess the impact of the fully operational HIS is often referred to as a *summative evaluation*. These types of evaluations are usually primarily quantitative and based on study designs intended to assess causality—randomized controlled trials (RCTs) and interrupted time series studies. In summative evaluations, the evaluators are usually more independent of the developers. These studies are usually at stage four and sometimes stage five in box 24.1.

Quantitative Evaluations

The main focus of this chapter is quantitative evaluation. In these studies, the goal is to assess the impact of using an HIS on some measurable endpoints. These could, for example, be the number of errors in drug prescribing or number of missed antenatal clinic appointments. Quantitative evaluations tend to be the best known and most commonly reported in the literature. They can provide rigorous evidence of the impact of the system on care, but may be less effective at explaining the underlying processes and sociotechnical issues that explain results (a key role for qualitative research studies). A range of study types include the following.

- *Demonstration studies.* These assess whether a system behaves as expected based on its design.
- *Comparative studies.* These include intervention and control groups to allow comparison of the intervention and usual care. These are the most common type in the literature.

Control Groups

Scientific studies of the impact of an intervention require a comparison group that stays the same when the intervention is introduced. There are many types of controls:

- historical
- contemporary
- matched
- randomized
- time series.

One of the simplest designs is a “before and after” study of clinical data collection and use, and is carried out when an HIS is implemented. If there are any changes measured in process and outcomes, these may be due to the new system. Unfortunately, these changes may be, and often are, due to unrelated changes in the health system, staffing, other technologies, and so on. Such *historical controls* do not work well for studies of complex interventions, particularly in LMIC where many changes may be happening at once. *Contemporary controls* are other sites, subjects, or groups that are compared to the intervention group. The approach can avoid the biases due to unrelated changes in the health system seen in historical controls, but unfortunately can introduce a different set of potential biases. The other sites may be different from the intervention sites, and therefore not provide a valid comparison. *Matched controls* are selected by statistical comparison with the intervention sites to be similar on key variables such as patient numbers, age, gender, and severity of illness, as well as characteristics like infrastructure, power, networking, and staff IT skills. This approach reduces the risk of bias, but finding good matches is an imprecise process (see chapter 4 for more in-depth discussion of these issues).

Causality: Determining Whether a Health Information System is Responsible for a Change Measured

As noted earlier, when an HIS is implemented, there are likely to be other changes in the health system that may be responsible for any observed differences. The best accepted way to show that changes are due to the HIS is to carry out an RCT. The key features of an RCT are the presence of one or more control groups that don’t have the system and

the randomization in the selection of these groups. If control subjects or groups are selected by nonrandom methods, it is very difficult to avoid baseline differences between the groups that can lead to bias. In typical RCTs for testing medications, there is an additional feature—double blinding. This means ensuring that neither the patient nor the doctor is aware of whether the patient is receiving the drug or the placebo. This reduces any bias in assessment of the patient. While such an approach could be valuable in studying HIS, it is rarely feasible to hide the use of the HIS.

There are other ways in which it can be possible to attribute causality to an HIS. One example is an *Interrupted Time Series*. In this case a group is studied (e.g., a group of patients) and after a certain period (period 1), the HIS is implemented for a while (period 2), then deactivated (period 3). This may be repeated multiple times. If there are improvements in the study group that are only seen with the HIS, then disappear when it is deactivated, it is likely they are due to the HIS, not other unrelated changes. An example is a decision-support system for chest pain patients incorporated into an ECG machine. In one study, the decision support was switched on and off multiple times [3]. A variant of this approach is studying the roll-out of an HIS such as an EHR to many sites over a period of time in a region or country by examining data submitted to a central site such as a national reporting system (e.g., DHIS2; see chapter 41 in book supplement). If site implementations are carried out on multiple known dates, then any changes in submitted data that occur *consistently* after the HIS is implemented are likely to be due to the system. Other methods for showing causality rely on “natural experiments” when the implementation of the HIS occurs in some site, but not others, due to factors unrelated to the local health system.

Evaluation Matrix: Stakeholders and Possible Evaluation Questions

Table 24.1 breaks down the different evaluation questions and types by the stage of HIS development and implementation and key stakeholders. It is intended to give examples of what type of evaluation study may be useful for a specific project.

Utilization-Focused Evaluation

An important innovation for studies in LMIC is utilization-focused evaluation [4]. This is an approach to ensuring that evaluations are of relevance and use to key stakeholders in the organization implementing the HIS, and are not just focused on questions of interest to, for example, academics and policy makers. The method employs standard stages: (1) identify key stakeholders; (2) determine their evaluation questions; (3) design studies to answer those questions; (4) feedback results to stakeholders. A potential disadvantage of the approach is that involvement of stakeholders may be perceived to bias the study. This may be addressed by involvement of external researchers in the design,

Table 24.1

Matrix of example questions for different stakeholders at each evaluation stage (items may apply to more than one stage)

Stakeholder	Requirements gathering	Design and development	Initial deployment	Scale-up	Long-term use
Clinician	What is the clinical problem? Clinical workflow, quality of care, efficiency of care	User interface, mapping of key activities and workflow, managing risk of errors or failure	Stability, speed, efficiency, security, adaptability, data quality, learnability	Quality of care, reducing errors, effect on speed of care delivery, ease of reporting and administration, evidence of clinical impact	Stability, extensibility, interoperability, maintainability, access to data for clinical care and research
Patient	Access to care, support in care, information, confidentiality	Assessment of information needs, user interface design and testing	Ease of use, quality of explanations, accessibility	Able to cover a range of key issues, delivery of key information like lab results [18], positive feedback from other patients	Good long-term stability and support, growth to new clinical areas
Health service manager	Quality, costs, efficiency	Cost of development, quality of software, suitability of architecture for interoperability and scalability	Stability, speed of implementation, quality of training resources, usage levels and satisfaction, data quality, cost of initial deployment [14], staff job satisfaction	Efficient scale-up and roll-out, compliance with clinical guidelines, evidence of clinical impact, efficient and accurate reporting, data quality, ability to monitor clinics and care processes	Stability, good usage and satisfaction levels, adaptability, total cost of ownership [31]
Developer	Accurate requirements and use cases	Effective software development strategy correctly applied, software testing strategy, user interface testing strategy	Effective implementation plan, training of implementers, formative evaluation of system	Monitoring of roll-out, measurement of stability, usage rates, technical problems	Total cost of ownership, effectiveness of interoperability, ability to support and upgrade system in the field.
MOH	Quality of care, access to care, surveillance, planning	Stability and safety of software, extensibility, interoperability, local support	Implementation plan, usage and acceptability by staff, data quality, reporting quality, and speed	Ability to scale up, scale-up planning, effectiveness of support for national health strategy	Stability, security, extensibility, interoperability, total cost of ownership

implementation, and analysis of the study or following up positive findings in larger and more independent study, or both.

Biases in Evaluation Studies of Health Information Systems

As noted above and in chapter 5, there are a number of potential biases in evaluation of HISs. These discussed in detail in a textbook by Friedman and Wyatt [5] and summarized here. While it is difficult to completely eliminate biases, good design can minimize their effects (+ and – symbols refer to biases that tend to increase or decrease respectively the apparent effect of the HIS).

- *Volunteer effect* (+): recruiting subjects who are keen to use the system and not typical users can lead to better outcomes than those from typical users.
- *Assessment bias* (+/–): in a non-blinded study, participating clinicians may rate intervention patients as healthier or better in some way than the controls when there is no real difference. This can also cause a negative bias.
- *Placebo effect* (+): the presence of the system may change behavior due to user expectations and lead patients to report better outcomes.
- *Checklist effect* (+): structured recording of data on a form can improve data completeness and quality, and therefore improve the processes of care being studied, independent of use of the HIS.
- *Hawthorne effect* (+): monitoring users as part of the study makes them change behavior and potentially try harder to work with the HIS.
- *Carry-over effect* (–): users who have access to decision support for some patients and not others learn from the system and improve care of control patients.
- *Allocation and recruitment bias* (+): evaluators recruit users or patients who are likely to benefit from use of the system.
- *Secular trends* (+): the problem of historical controls; many other changes and improvements are going on at the same time as the study, making improvements hard to attribute to the HIS.

Qualitative Evaluations

These studies focus on the experience of the users of an HIS and assess issues like user satisfaction, ease of use, barriers to effective use of HIS, and social and cultural influences on adoption and use of systems. Clearly, these issues can be of high relevance in LMIC, which have a range of environments and limited experience with information technology. This information is valuable in assessing how users interact with the system and whether those interactions are positive or negative. A key benefit of such studies is understanding

how well the HIS fits with the workflow and activities of users and how the system design can be improved. User feedback usually provides important insights into the likelihood that the system will scale and also have a beneficial impact.

Kaplan and Maxwell describe the use of qualitative methods for health informatics in detail [6]. They suggest a qualitative approach can be helpful for the following research questions and situations:

- to determine what might be important to measure or why measured results are as they are; if the subject of study cannot be measured easily.
- to understand not only what happened or what people are responding to, but why; to understand how people think or feel about something and why they think that way, what their perspectives and situations are and how those influence what is happening; to understand and explore what a technology (e.g., a newborn nursery telemonitoring system) or practice (e.g., using a computer to access health information) means to people.
- to investigate the influence of social, organizational, and cultural context on the area of study, and vice versa.
- to examine causal processes, and not simply what causal relationships exist.
- to study processes as they develop and emerge, rather than in outcomes or impacts; for example, to investigate the development process for the application under study in parallel with that process so that you can improve the application development as it progresses.

Methods used in qualitative studies of HIS include semi-structured questionnaires to elicit user views and experiences, focus groups, and observation of users interacting with systems under real-world or experimental conditions. Rose et al. describe the use of “qualitative studies to improve the usability” of an EHR, emphasizing the importance of assessing user needs and experiences with a “results manager” component of an EHR. They used task analysis and focus groups and combined the results to make recommendations for improving the design of the results manager [7]. In another example, Zakane et al. used semi-structured interviews to study the “needs of and attitudes towards a computerized clinical decision-support system in rural Burkina Faso.” They found interest in and support for an HIS and decision-support systems to improve maternal and child health, but concerns that the systems would be difficult to learn and might increase work load [8].

These methods also help researchers to understand the meaning and context of items being studied and the underlying dynamics of the processes and the way they impact users. Chapter 6 discusses some of the consequences of implementing health interventions in LMIC and the importance of studying the broader consequences and potential risks. Related to these ideas, Farach and colleagues used semi-structured interviews conducted

mainly by Skype to assess the impact of HIS initiatives in addressing the health needs of underserved populations in Latin America and the Caribbean. They reported that HISs can benefit the health needs of such populations [9].

Some researchers go as far as arguing that quantitative, sometimes called “objectivist” studies cannot capture the important sociotechnical interactions that are the real determinants of success and failure of HISs, and therefore are not valid without qualitative studies [10]. There are many examples of technically sound HIS projects that failed due to sociocultural issues. Conversely, there are many projects that failed due to poor design, development, and implementation of software or infrastructure weaknesses. As this chapter should make clear, there are a range of factors that need to be considered in effective evaluation studies, and qualitative research methods are a key element in the mix. While there is not space here to provide an in-depth discussion of the principles and techniques of qualitative research, readers are encouraged to consult the reading list and also consider collaborating with qualitative researchers in their studies.

Economic Evaluation

One area of HIS deployment in LMIC that currently has very little evaluation evidence is the cost of these systems. There are several key principles that are important in assessing the real costs of an HIS and comparing those with evidence of benefits. Readers are encouraged to explore this important topic in more detail (e.g., in references [11] and [12]).

Accurately measuring the true cost of an HIS requires assessing the following items required for *initial implementation*:

- hardware and IT costs
- improvements needed in infrastructure, such as building work, power, networking, and air conditioning
- software development or purchase, including all adaptation and improvement costs
- implementation costs, including staff time for training and reduced productivity during changeover
- training of users and support staff
- IT support staff.

In addition, the costs of *running the system over several years* must be measured. These include:

- ongoing staff costs for IT support, training, data management, analysis, and so on
- replacement hardware such as servers, clients, and networking
- supplies and monthly costs such as printing, network subscriptions, and power

- ongoing software costs, which could include licenses, essential improvements, upgrades, and security fixes.

Combining these items creates the *total cost of ownership*. This figure is critical in assessing which HIS projects make most sense, since initial costs may give a misleading impression of the longer-term costs. This is especially true with proprietary software with high license fees and projects dependent on reliable network connections, which may be expensive (and possibly unreliable despite the cost).

More advanced economic evaluations take into account the impact of the HIS as well as the cost. These include *cost-effectiveness* analysis, where the total cost of ownership is combined with measured benefits of the HIS to show the cost of a specific, measured improvement per specific health outcome, patient, or health facility. *Cost-benefit* analysis is similar, except that it considers an array of benefits and assigns a financial value to each—a more difficult and controversial approach, especially in LMIC. Both techniques rely on an accurate evaluation of the actual benefits of the HIS, so must be combined with rigorous outcome studies.

Two examples of costing studies illustrate some of these key principles. Saronga and colleagues examined the cost of implementing a decision-support system for maternal care in Tanzania. They detail the many items that need to be costed, including vehicles and transport, computers and IT hardware, furniture, buildings, software, personnel, training, and supplies and communication. These were grouped into “installation and operation costs; capital and recurrent costs; and fixed and variable costs.” The total cost reported was \$185,927, with 77% incurred in the installation phase. Training and software were the largest cost items, at 33% and 32% of total cost, respectively [13].

A study by Driessen and colleagues looked at the potential cost savings from implementing an EHR system in a hospital in Lilongwe, Malawi [14]. They costed both the implementation of the EHR and the running cost over 3–5 years. They also modeled the costs that could be saved by improving efficiency of processes, such as the labeling and tracking of laboratory samples. The hospital typically had substantial losses due to discarded samples and repeating tests due to poor labeling. This problem was addressed by the creation of bar-coded ID labels and cards from the EHR system. An actual prospective evaluation is required to assess the true costs and savings.

Examples of Key Studies in LMIC

There are an increasing number of studies of HISs in LMIC being published over the last decade. An early systematic review of the literature in 2010 showed initial evidence of benefits of these systems in a number of areas [15]. These benefits have been largely confirmed by more recent and rigorous studies, some of which are enumerated below:

- tracking patients through treatment initiation and monitoring adherence, then detecting those at risk for loss to follow-up (e.g., text messages to increase antenatal clinic adherence [16] and to encourage HIV-positive patients to return for treatment [17])
- decreasing time to create administrative reports (many unpublished studies)
- tools to label or register samples and patients (such as labeling and tracking of samples in a hospital in Malawi [14])
- collection of clinical or research data using personal digital assistants (now confirmed in a number of mHealth projects)
- reduction in errors in laboratory and medication data [18,19]
- reminding patients of health care actions (such as text-messaging reminders for HIV treatment [20,21], although there are also some negative studies of these approaches)

More recent systematic reviews have examined the impact of mHealth systems on health service delivery processes [22], and on care of patients with HIV [23]. They have shown progress in certain areas such as text-messaging reminders to patients to take medications for HIV, but emphasized the need for more and better studies. Based on the experience of this study, Free et al. also note the importance of clearer descriptions of the underlying interventions to better understand what was done and how to reproduce it.

In 2011, there was a meeting in Bellagio, Italy, to encourage more and better evaluation studies in LMIC. This led to the development of a call to action for eHealth evaluation and a set of nine principles to do so (see the recommendations for further reading at the end of this chapter).

Data Quality Improvement

Good data quality is absolutely fundamental to achieving positive benefits from HISs. Worldwide, all HISs have challenges with data quality, though these are often greater in LMIC.

Data quality is primarily defined by the following three metrics.

- Data completeness (is there missing data?)
- Data timeliness (is data entered in time for required uses?)
- Data accuracy (is data a true reflection of the real-world items being measured?)

A useful principle is to determine if the data is “fit for purpose.” This metric will differ depending on the type of use. So, for example, data for clinical care or research needs to be highly accurate, but timeliness requirements may vary. It may be acceptable for research data to be entered sometime after the patient is seen, but clinical data needs to be available as soon as possible, especially in a hospital setting. While it is unlikely that HIS

projects can completely eliminate poor data, there are several strategies that should be applied to improve it:

1. ensure that staff collecting and entering data can see how it is used and, in general, get benefit from their own data. When clinical staff and other health care workers enter data directly, they should be able to see what they entered and summaries or reports created from it. Many public health projects in LMIC have a “one-way ticket” with staff who are required to collect data, for example for a monthly report on HIV patient care, sending it off on forms to be used at district or national levels and never seeing the results [24].
2. provide training to staff collecting and entering data to help them maintain high accuracy. This should be supplemented with tools to measure data quality and opportunities for staff to recommend ways to improve data quality.
3. ensure that the HIS is well designed and tested, possesses a good user interface and clear methods for coding and export of data, and runs on stable hardware and networks. For direct data entry by clinical staff, the system must be available during all clinic hours. Significant down time is highly disruptive and will result in loss of confidence in the system and likely reduced use.
4. ensure that a unique ID is issued for each patient and used for all entries in the EHR. This will reduce errors, for example, mis-entry of data into the wrong record, duplicate entries that are not easily found by clinicians, and poor naming of lab samples leading to lost results.
5. automate uploading data from other systems where possible. This is particularly valuable for laboratory data. Many EHR systems can upload lab data using open interoperability standards like HL7.
6. resist the temptation to collect more than the essential data needed for clinical care, or if additional data is required for example for research, add extra staff and resources.

Most importantly, data quality must be measured regularly—ideally, daily, and action taken promptly to fix problems. Several studies have shown significant improvements in data quality for care of HIV, multidrug-resistant TB, and maternal health patients with the implementation of HISs in LMIC, via the use of better tools and training [18, 25–27].

Example: Real-Time Assessment of System Performance and Data Quality

An example of monitoring and evaluation of system performance and data quality is being implemented in Rwanda as part of an evaluation of the impact of an EHR system (OpenMRS) that has been rolled out to over 300 clinic sites by the Rwandan Ministry of Health. Many sites have limited or unstable power, and Internet access is variable and

sometimes completely unavailable. To ensure intervention sites are performing well, each will have the following metrics measured:

1. EHR system uptime and downtime, both frequency and total duration
2. EHR data completeness for key variables
3. EHR system usage, including viewing of key pages and data entry into forms by staff.

These metrics are being measured with a custom software module, and results will be transmitted to a central server on a daily basis (via the mobile network in many sites). In addition, periodic site visits will be made to assess data quality in the EHR versus paper records or other data sources. The study will also survey users on their experience using the EHR and what improvements may be required. This will be supplemented by a qualitative research study interviewing a subset of users and observing them working with the HIS. This study can be classed as a formative evaluation of system performance (or sometimes a Process Evaluation) and will be followed by actions to improve sites where performance falls below certain thresholds on one or more metrics.

The Learning Health System

A range of new initiatives are bringing together many of the key initiatives described in this chapter and more broadly in the book to create a system for continuous measurement and quality improvement in health systems. Entitled the Learning Health System, part A (figure 24.1) combines the processes of data collection and analysis with the development of new knowledge. This part of the process may lead to publication of clinical insights. Part B describes the process by which the insights into clinical care and workflow are turned into guidelines and clinical decision-support tools (or other health care improvements) to support improved quality of care.

Critical to these processes are the feedback loops that support clinical care. Loop 1 is use of clinical data for direct care, an essential part of ensuring clinician buy-in and oversight of data quality. Loop 2 is the development of knowledge based on analysis of multiple records (Part B above). Loop 3 is the publication of new knowledge that informs broader health care initiatives. By combining all three processes based on routine clinical data collection, resources can be shared for data-quality control processes.

Conclusion and Future Directions

Evaluation studies are intimately linked to methods and processes to improve the quality of HIS projects. Working through the different stages of evaluation and addressing the needs of key stakeholders should ensure that appropriately targeted, well-designed and programmed HISs are created that are also easy to use and viewed as helpful by users.

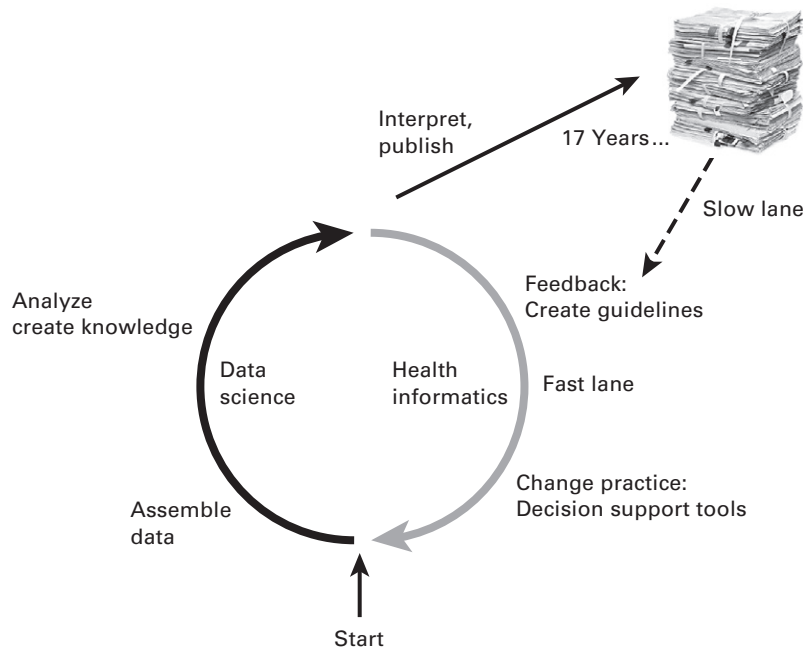


Figure 24.1
The learning health system (adapted from Friedman [28]).

A high-quality HIS of this sort will not guarantee success in all projects, since the complexities of different organizations and cultures may lead to unanticipated problems. However, most big failures of HISs are attributable to major errors in requirements gathering, design, programming, or implementation—known and preventable problems [29]. Frequently, systems are not complete or not designed for the purpose and workflow for which they are being used, or they are difficult to use and lack user testing. A particular challenge in LMIC are HISs that require much better infrastructure than is available, including power, networking, Internet, and effective IT support. The process of monitoring daily usage of an HIS, including uptime, page views, and data completeness, goes a long way to ensure that failures are detected early—offering opportunities to correct the problem or in some cases cancel the project and go back to the drawing board.

Another strategy for successful system implementation is to combine “bottom-up” and “top-down” approaches. The bottom-up approach ensures that HIS projects address clinical needs and problems, focus on local ownership and priorities, and prioritize data quality. These approaches help to ensure the project works well in the original site. The top-down approach includes planning for wide use, a horizontal approach to health care delivery where systems can cover a whole range of diseases and clinical problems, along with a

focus on creating a core dataset and supporting open standards and interoperability. Both approaches should be married, along with a focus on evaluation and evidence-based decision making.

Finding the resources, expertise, and time to carry out effective evaluation of an HIS remains a concern for many organizations, but, as described here, there are many different approaches that can be applied. In addition, even small, formative evaluations can be very useful, both for the implementing organization and as lessons for the wider community [30]. The alternative cost of avoiding evaluation can be very high—large-scale projects that fail in expensive and embarrassing ways due to elementary and preventable errors.

Questions for Discussion

- Choose a project you are familiar with and critically examine what tools are in place to monitor and improve data quality. What else should be done to maximize data quality?
- Choose a specific evaluation question and list all the stakeholders likely to be interested in it. At what stage or stages in the life cycle of a health information system is it most relevant to carry out an evaluation?
- Taking the example from the previous question, what biases are likely to be important in the evaluation study you have in mind?
- A randomized controlled trial performed a decade ago showed that a web-based health information system can improve the timeliness of, access to, and quality of laboratory data. Is that result likely to be relevant today for a project planning to use an mHealth application to support laboratory reporting? What is different, and what is likely to be similar?

Further Reading

The following articles and books provide additional depth for readers interested in carrying out evaluation studies.

- Friedman CP, Wyatt J. *Evaluation Methods in Biomedical Informatics*. Springer Science and Business Media; 2006. Textbook providing a comprehensive overview of the subject and how to perform evaluations studies, third edition due to be published 2015.
- Bellagio eHealth Evaluation Group. 2011. Call to Action on Global eHealth Evaluation. Consensus Statement of the WHO Global eHealth Evaluation Meeting, Bellagio, Italy, September.
https://www.ghdonline.org/uploads/The_Bellagio_eHealth_Evaluation_Call_to_Action-Release.docx.

- Were MC, Nyandiko WM, Huang KT, Slaven JE, Shen C, Tierney WM, Vreeman RC. 2013. Computer-generated reminders and quality of pediatric HIV care in a resource-limited setting. *Pediatrics* 131(3): e789–e796.
A key study showing use of RCT methodology to demonstrate impact of eHealth system on quality of HIV care in Kenya.
- Blaya JA, Shin SS, Yagui M, Contreras C, Cegielski P, Yale G, Suarez C, et al. 2014. Reducing communication delays and improving quality of care with a tuberculosis laboratory information system in resource poor environments: A cluster randomized controlled trial. *PLoS ONE* 9(4): e90110.
A key study showing use of RCT methodology to demonstrate impact of eHealth system on access to critical laboratory data to improve management of drug resistant TB in Peru.
- Nykanen P, Brender J, Talmon J, de Keizer N, Rigby M, Beuscart-Zephir MC, et al. 2011. Guideline for good evaluation practice in health informatics (GEP-HI). *Int J Med Inform* 80(12): 815–827.
A valuable framework for carrying out and documenting evaluation studies in Health Informatics.
- Aqil A, Lippeveld T, Hozumi D. 2009. PRISM framework: A paradigm shift for designing, strengthening and evaluating routine health information systems. *Health Policy Plan* 24(3): 217–228.
This article describes a classic evaluation framework for health information systems with a focus on LMIC and patient-level data.

References

1. Fraser HSF, Blaya J, Choi SS, Bonilla C, and Jazayeri D. 2006. Evaluating the impact and costs of deploying an electronic medical record system to support TB treatment in Peru. *AMIA Annu Symp Proc* 2006: 264–268.
2. Stead WW, Haynes RB, Fuller S, Friedman CP, Travis LE, Beck JR, et al. 1994. Designing medical informatics research and library—resource projects to increase what is learned. *J Am Med Inform Assoc* 1(1): 28.
3. Selker, HP, Beshansky JR, Griffith JL, Aufderheide TP, Ballin DS, Bernard SA, et al. Use of the acute cardiac ischemia time-insensitive predictive instrument (ACI-TIPI) to assist with triage of patients with chest pain or other symptoms suggestive of acute cardiac ischemia: A multicenter, controlled clinical trial. *Ann Intern Med* 129(11): 845–855.
4. Patton MQ. *Utilization-Focused Evaluation*. Sage Publications; 2008.
5. Friedman CP, Wyatt J. *Evaluation Methods in Biomedical Informatics*. Springer Science and Business Media; 2006.

6. Kaplan B, Maxwell JA. Qualitative research methods for evaluating computer information systems. In Anderson JG, Aydin CE, eds. *Evaluating the Organizational Impact of Healthcare Information Systems*. 2nd ed. New York: Springer; 2005: 30–55.
7. Rose AF, Schnipper JL, Park ER, Poon EG, Li Q, Middleton B. 2005. Using qualitative studies to improve the usability of an EMR. *J Biomed Inform* 38(1): 51–60.
8. Zakane SA, Gustafsson LL, Tomson G, Loukanova S, Sié A, Nasiell J, et al. 2014. Guidelines for maternal and neonatal point of care: Needs of and attitudes toward a computerized clinical decision support system in rural Burkina Faso. *Int J Med Inform* 83(6): 459–469.
9. Farach N, Faba G, Julian S, Mejía F, Cabieses B, D'Agostino M, Cortinois AA. 2015. Stories from the field: The use of information and communication technologies to address the health needs of underserved populations in Latin America and the Caribbean. *JMIR Public Health and Surveillance* 1(1): E1.
10. Greenhalgh T, Russell J. 2010. Why do evaluations of eHealth programs fail? An alternative set of guiding principles. *PLoS Med* 7(11): E1000360. doi:10.1371/journal.pmed.1000360.
11. Levin H, McEwan P. 2001. *Cost-effectiveness analysis: Methods and applications*. Thousand Oaks, CA: Sage Publications.
12. Tan-Torres Edejer T, Baltussen R, Adam T, Hutubessy R, Acharya A, Evans DB, et al., eds. 2012. Making choices in health: WHO guide to cost-effectiveness analysis. Geneva, Switzerland: World Health Organization. http://www.who.int/choice/publications/p_2003_generalized_cea.pdf?ua=1.
13. Saronga HP, Dalaba MA, Dong H, Leshabari M, Sauerborn R, Sukums F, et al. 2015. Cost of installing and operating an electronic clinical decision support system for maternal health care: Case of Tanzania rural primary health centers. *BMC Health Serv Res* 15(1): 132.
14. Driessen J, Cioffi M, Alide N, Landis-Lewis Z, Gamadzi G, Gadabu OJ, et al. 2013. Modeling return on investment for an electronic medical record system in Lilongwe, Malawi. *J Am Med Inform Assoc* 20(4): 743–748.
15. Fraser HSF, Blaya J. 2010. Implementing medical information systems in developing countries, what works and what doesn't. *Proc AMIA Symp* 2010: 232–236.
16. Lund S, Nielsen BB, Hemed M, Boas IM, Said A, Said K, et al. 2014. Mobile phones improve antenatal care attendance in Zanzibar: A cluster randomized controlled trial. *BMC Pregnancy Childbirth* 14(1): 29.
17. Siedner MJ, Lankowski AJ, Kanyesigye M, Bwana MB, Haberer JE, Bangsberg DR. 2015. A combination SMS and transportation reimbursement intervention to improve

HIV care following abnormal CD4 test results in rural Uganda: A prospective observational cohort study. *BMC Med* 13(1): 160.

18. Blaya JA, Shin SS, Yagui M, Contreras C, Cegielski P, Yale G, et al. 2014. Reducing communication delays and improving quality of care with a tuberculosis laboratory information system in resource poor environments: A cluster randomized controlled trial. *PLoS One* 9(4): E90110.

19. Choi S, Jazayeri D, Mitnick C, Chalco K, Pachao F, Bayona J, et al. 2004. A web-based nurse order entry system for multidrug-resistant tuberculosis patients in Peru. *Proc. Medinfo* 11: 202–206.

20. Lester RT, Ritvo P, Mills EJ, Kariri A, Karanja S, Chung MH, et al. 2010. Effects of a mobile phone short message service on antiretroviral treatment adherence in Kenya (WelTel Kenya1): A randomised trial. *Lancet* 376(9755): 1838–1845.

21. Pop-Eleches C, Thirumurthy H, Habyarimana JP, Zivin JG, Goldstein MP, De Walque D, et al. 2011. Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: A randomized controlled trial of text message reminders. *AIDS* 25(6): 825.

22. Free C, Phillips G, Watson L, Galli L, Felix L, Edwards P, et al. 2013. The effectiveness of mobile-health technologies to improve health care service delivery processes: A systematic review and meta-analysis. *PLoS Med* 10(1): E1001363.

23. Catalani C., Philbrick W, Fraser H, Mechael P, Israelski DM. 2013. mHealth for HIV treatment and prevention: A systematic review of the literature. *Open AIDS J* 7: 17–41.

24. Mate KS, Bennett B, Mphatswe W, Barker P, Rollins N. 2009. Challenges for routine health system data management in a large public programme to prevent mother-to-child HIV transmission in South Africa. *PLoS One* 4(5): E5483.

25. Amoroso C, Akimana B, Wise B, Fraser HSF. 2010. Using electronic medical records for HIV Care in rural Rwanda. *Stud Health Technol Inform* 160: 337–341.

26. Blaya JA, Cohen T, Rodríguez P, Kim J, Fraser HSF. 2009. Personal digital assistants to collect tuberculosis bacteriology data in Peru reduce delays, errors, and workload, and are acceptable to users: Cluster randomized controlled trial. *Int J Infect Dis* 13(3): 410–418.

27. Haskew J, Rø G, Saito K, Turner K, Odhiambo G, Wamae A, et al. 2015. Implementation of a cloud-based electronic medical record for maternal and child health in rural Kenya. *Int J Med Inform* 84(5): 349–354.

28. Learning Healthcare Project. <http://www.learninghealthcareproject.org/section/background/learning-healthcare-system>. Accessed 28/11/2016

29. Coiera E. 2013. Why e-health is so hard. *Med J Aust* 198(4): 178–179.
30. Fraser HSF, Blaya J, Choi SS, Bonilla C, and Jazayeri D. 2006. Evaluating the impact and costs of deploying an electronic medical record system to support TB treatment in Peru. *AMIA Annu Symp Proc* 2006: 264–268.
31. Dalaba MA, Akweongo P, Aborigo RA, Saronga HP, Williams J, Blank A, et al. 2015. Cost-effectiveness of clinical decision support system in improving maternal health care in Ghana. *PLoS One* 10(5): E0125920. doi:10.1371/journal.pone.0125920.