Evaluation of Electronic Health Record Systems

C. Nøhr
Virtual Centre for Health Informatics, Department of Development and Planning, Aalborg University, Denmark

Summary
Objectives: The objective of this article is to present an overview of the various considerations to be made prior to evaluating electronic health record (EHR) systems.
Methods: From the methodological literature, a number of themes for decisions are presented and related to the contemporary EHR situation. Special attention is paid to a number of important methodological themes.
Results: Definitive checklists for evaluation of EHR systems cannot be recommended, but seven key steps are listed to guide the design of evaluation projects.
Conclusion: It is concluded that the issues presented are not completely exhausted and the seven key steps might have to include iterative loops because of interdependencies between some of the steps.

Introduction
Among the Grimm brothers’ numerous fairy tales, many remember the one called “Sweet Porridge”. It is about a poor little girl who lived alone with her mother, and they no longer had anything to eat. The little girl went into the forest where she met an aged woman who had a little pot. When she said, “Cook, little pot, cook,” it would cook good, sweet porridge, and when she said, “Stop, little pot,” it ceased to cook. The little girl took the pot home to her mother. One day, when the mother was on her own and wanted some porridge she had the pot cook some, but she did not know how to stop it. The porridge rose over the edge, and still it cooked until the kitchen and the whole house were full, and then it ran out into the street and filled the whole village. At first, the other inhabitants appreciated the unanticipated meal but eventually it nearly choked them because the porridge was everywhere.

There is a rising tide of evaluation approaching – many health informaticians are already in “porridge” to the waist. To avoid being choked to death by the complexity of evaluating they focus on methodologies that they are familiar with and define these as “the best way to evaluate”. Some evaluators move into niches and restrict their evaluations to subsets of IT-applications or specific systems; other health informaticians declare themselves in opposition to undertake evaluation. In spite of the discrepancy, evaluation is increasingly institutionalized: in national evaluation centers, technology assessment boards, etc. Special working groups on evaluation have been established under the European Federation of Medical Informatics (EFMI EVAL) as well as under the International Medical Informatics Association (IMIA WG15). The aim of this article is to present an overview of recent methodological discussions, and relate these discussions to the evaluation of electronic health record (EHR) systems. The aim is not to draw attention to one specific superior method for evaluating EHR systems, or present a detailed checklist for such evaluations. Furthermore, the article does not present a review of all the evaluation projects that have been reported in the scientific literature. A web-based inventory of evaluation studies in health informatics has been set up at the University of Medical Informatics and Technology Tirol by the Research Group Assessment of Health Information Systems [1]. This inventory contains 1035 evaluation studies published between 1982 and 2002 and is an excellent source of inspiration to the specific design of a local evaluation study.

The complexity of the systems and the environment in which they function make evaluations of EHRs complicated. Therefore, a number of delimiting decisions have to be made in order to design a practical evaluation study. These decisions fall in five categories

• System to evaluate
• Technical design or functionality
• Life cycle stage
• Stakeholder viewpoint
• Methodology

Not all categories are complicated to decide on – it is rather a question of making up ones mind. Some of the categories are unfolded and elaborated on in the following.
What is an EHR System?

In the literature there are numerous acronyms for systems handling patient data: Electronic medical record (EMR), electronic patient record (EPR), computerized medical record (CMR), computer-based patient record (CPR), and electronic health record (EHR).

There are only minor differences in the meanings depending on the defining country of origin, health sector, professional discipline, and period of time. The acronym EHR is preferred in this article as the term health refers to a person’s vital data independent of the specific periods of being a patient.

There have been many formal definitions of these terms, but they display more similarities than differences with respect to the purpose, functions and goals of electronic records [2].

The International Standard Organization (ISO) has developed an internationally agreed definition of EHR: “a repository of information regarding the health of a subject of care, in computer processable form”. Here the EHR is defined in terms of its structure, but it is internationally broad and encompasses most of the EHR systems currently used.

These very broad definitions make it difficult to determine the exact object of an evaluation. EHR systems that encompass all the functionalities of the above definition would also be impossible to manage and implement at once. Therefore EHR systems are normally subdivided into components or modules each handling one or more functionalities. The most common components are:

**Clinical Documentation**
Handles progress notes either as free text directly entered into the system or via predefined structured notes. Voice recognition systems are also seen as data entry method.

**Physician Order Entry (POE)**
Used for ordering diagnostic test and medication in a standardized and formalized way. Some systems also check for drug interactions and alert for patient allergy.

**Booking service**
Allows patients and clinicians to book appointments.

**Communication/Messaging**
Facilitates communication between hospitals, General Practitioners, pharmacies, and laboratories.

**Results Management**
Abnormal results warning, trending/graphing.

**Charge Capture/Billing**
Coding interventions (this module varies a great deal according to the organization and financing of the health sector).

**Decision Support**
There are a lot of decision support modules in use, but they are rarely used outside the organization where they are developed.

**Clinical Practice Guidelines**
Module to manage and maintain clinical guidelines or national reference programs – sometimes categorized as decision support.

**Disease Management**
Management of chronic diseases, like diabetes, etc.

**Management of security issues**
All EHR systems will have special facilities to manage authentication and authorization of user access according to national legislation.

In practice there will be difference between the modules in which a specific functionality is positioned i.e. the above list can neither be regarded as exclusive nor exhaustive. This is also true for the specific way a work task is executed in practice. E.g. one department can use the booking module to book an x-ray for a patient where another department will use facilities in the CPOE module. This makes it difficult to generate evaluation results that can be generalized across many hospitals.

The development of EHR systems can be traced back to the 1960’s but it was not until the 1980’s that implementation really began. Only a few countries in Europe use EHR systems on a large scale [2]. More countries have now developed national strategies to implement EHR systems for managing and communicating patient information [3-6].

Bringing EHR systems from the developmental stage into daily routine makes the evaluation of these systems topical. The concern for evaluation comes from many sources: clinicians want to know whether the system is of any benefit to their work procedure and will improve quality to the benefit of their patients, managers of the health care organizations wish to assure that the technology implemented contributes to production effectiveness, developers and vendors are interested in prioritizing areas for improving their product, politicians or the financing party want to ascertain that the resources are optimally spent, and the patients want to be reassured that the system confidentiality complies with legislation. Of course the different stakeholders have mutual interests in many areas.

Implementing EHR systems with a multitude of functionalities is a non-trivial issue that can stretch over a long period of time. Some hospitals have spent several years (from 3 to 12 years) implementing CPOE systems [7], which also means that significant end results appear much later. From the above it should be clear that developing and implementing EHR systems are tremendously complex tasks and therefore even more complicated to evaluate unambiguously. At the entry of a concrete evaluation project at least three decisions must be made:

1. What to evaluate: functionality or technical design,
2. At what stage of the lifecycle will the evaluation take place: development or implementation or …
3. Which stakeholder’s viewpoints should be pursued.
These decisions will establish a foundation for determining the methodological base for the specific evaluation.

Methodology for Evaluating EHR Systems

While a method is a codified series of steps taken to complete a certain task or to reach a certain objective, a methodology is the study of the body of methods employed in a discipline i.e. evaluation projects. As there is no unique way to do evaluations and no generic logical structure that will ensure a unique method of choice, it is crucial to consider a few methodology issues prior to selecting a specific method. A set of opposing methodological issues are presented in the following.

• Objectivist versus subjectivist approach
• Qualitative versus quantitative methods
• Reductionist versus holistic approach
• Summative versus formative evaluation

Here the issues are presented as dichotomies, but in the design of real evaluation projects they are often applied more pragmatically. The main point, however, is that it is important to carefully consider each step in an evaluation project (a stepwise approach is outlined in the end of this paper).

The *objectivist* approach is based on measurements. It is assumed that it is possible, on a rational basis, to obtain measures that capture the knowledge of the issues that are evaluated. Further it assumes that a general truth exists and that this truth is established through objectivity of the measurements. In most areas of natural science the objectivist approach is well suited to investigate biological issues such as organs and tissues. In health informatics evaluation typical measurements have been response time of an EHR system in the dialog with a user, number of paper forms reduced, numbers of blood samples handled on one day, etc.

The *subjectivist* approach is concerned with emotional, cultural or tacit phenomena. These issues are often studied by means of observation, interviews and/or open questionnaires. The subjectivist studies do not seek to prove a hypothesis or demonstrate a specific statement, but rather to generate insightful descriptions that will enhance understanding. In health informatics typical questions studied by a subjectivist approach have been: how do users perceive the usability of an EHR system? Who is actually using the CPOE system for medication? How is the job satisfaction among the clinical staff after introduction of an EHR system? Etc.

*Qualitative methods* are applied where the endpoints of the evaluation are descriptive - where it is impossible or desirable to put a figure on the items that are evaluated. A scale often used in qualitative studies is the nominal scale. It is used to classify different variables in relation to each other, e.g. sex or social status. Each category must be well defined i.e. exclusive: anything should have a definite class, and exhaustive: everything that will be classified must find a class to fit into. On this scale no quantitative information is conveyed and no ordering of the items is implied.

*Quantitative methods* are evidently applied where there is a recognized scale to measure the variables. Basically there are four different categories of quantitative scales:

- **Ordinal scale**: Used when an item is classified according to whether they have more or less of a characteristic. However the difference between values is not necessarily equal. This scale is often used in questionnaires where respondents are asked to indicate their opinion.
- **Interval scale**: a linear ranking scale – the same distance between units, but with no natural zero. An example could be the measurement of blood glucose in mmol/l.
- **Ratio scale**: like an interval scale, but with a true zero point. It is now meaningful to talk about the double value. Height and age are examples of ratio scale variables.

A *reductionist* approach assumes that the issues that are evaluated consist of components that can be taken apart for separate analysis or measurement and that the sum of knowledge of these components allow us to explain the whole.

The *holistic* approach assumes that the whole consists not only of the individual components, but also other factors, which in mutual dependency is critical in making the whole more than the sum of the parts. In a holistic perspective it is therefore impossible to evaluate an EHR system by analyzing the technical issues and the organizational issues separately, and conclude that 80% of the problems encountered were caused by organizational issues.

*Summative* evaluations are done to account for the expected or promised outcome, primarily designed to demonstrate the value of a system. It is evaluation activities with the purpose of contributing with a status of a project and answer questions like: did the system fulfill the promises in the contract? Is the funded money spent according to budget?

*Formative* evaluations focus on continuous improvements of the system. Development and implementation of EHR systems are inevitably compromising between the ideal and the practicable. The abundance of compromises and decisions are made on local premises and the formative evaluation is, in a systematical way, providing the information basis for these decisions.

In between formative and summative evaluations there is a third kind of evaluation – *monitoring*. To monitor is to watch or to keep track of something. It has become a very common activity in quality assurance activities within the health care sector. The purpose of many clinical databases is simply to be a tool for monitoring the quality or effectiveness of the clinical work through a set of performance indicators.
Monitoring as an evaluation activity is not reflected in the theoretical and methodological literature as summative and formative evaluations. Nevertheless, monitoring is increasingly seen accompanying, for instance, national planning and strategies for health information systems including EHR.

The evaluation methodologies in health informatics are addressed by a number of publications [8-13]. The textbooks by Anderson et al. [8] and van Gennip et al. [13] are both multi-authored collections of articles presenting a spectrum of methodological approaches complemented by a number of cases. The textbook by Friedman and Wyatt from 1997 [12] presents a focused and homogeneous picture of evaluation methodology. Among other fundamental evaluation methods concepts they introduce the distinction between objectivist and subjectivist approaches.

The text published by Brender 1997 [10] is a monograph and a slightly edited version of a European Doctorate Thesis. It contains a literature survey, development of a methodology for assessment of knowledge based systems, a full scale prospective case study and at last a synthesis of the dynamic assessment methodology. The text contains a lot of useful examples which can give inspiration to the selection of dedicated methods for evaluation purposes. It does not provide a cookbook on how to perform an evaluation project. On the contrary it is argued that it is not possible to make comprehensible cook-book evaluations.

Brender has recently published a new book on evaluation methodologies [11]. This book is an elaboration of the ideas presented in [10]. The ideas are communicated in a handbook manner which enables the reader to plan and perform evaluation projects in a systematic and explicit way. Again it must be emphasized that this is a handbook where the reader can get ideas and references and not a cookbook containing evaluation recipes.

**Evaluation in Organizations**

Evaluations of EHR systems are always done in organizations unless it is a simple (laboratory) test of a technical detail. Hence the relation between technology and organization becomes important. Basically there are two different theories on how technology and organization relate, and the key questions are whether technology is regarded as an exogenous and stable entity or an endogenous variable entity and whether the technology is disseminated in diffusion processes or translation processes.

**Technology as an Exogenous Entity**

In the contemporary discussion the dominating notion is that technology and organization are separated in two contingent variables, and these two variables must somehow fit each other if the organization is to perform efficiently. This is also seen in the literature that addresses organizational change and change management [14;15]. A classic model that builds on the idea of interdependency between a number of variables is Leavitt’s model for organizational change [16]. In his four-variable conception of organizations (1) task refers to “the organizations raison d’être”, which is the production of goods and services including the many related subtasks. (2) Actors refers chiefly to people. (3) Structure refers to the system of communication, authority and workflow. And (4) technology which refers to machines, hardware as well as software. The variables are highly interdependent as shown in figure 1. The arrowheads indicate that change in any one variable results in compensatory change in others. In analyzing organizational change Leavitt assumes that it is one or more of these variables which are to be changed. Often the ultimate aim is to influence the task variable.

When technology and organization are separated into two variables which mutually impact each other the analysis and discussion about change tends to create a competition between the two entities. The technical world will accuse the users of being resistant, badly educated and ignorant, and the users in the organization will maintain that the hardware is too impractical and the software not built to a complex clinical world. As a result it will be impossible to see how the technological and the organizational are fundamentally entangled and mutually change each other, thereby producing unexpected results.

When the technology is seen as an exogenous entity the diffusion process becomes linear where the technology remains the same during the implementation process.
Technology as an Endogenous Entity

An alternative view exists in the notion that technology is an endogenous entity in an organization and vice versa. It is not a separate variable that impacts the organization – other things being equal. This view is expressed in organizational theoretical definitions of technology:

“…Most organization theorists have embraced the broader view that technology includes not only the hardware used in performing work, but also the skills and knowledge of workers, and even the characteristics of the objects on which work is performed.” [17].

When technology is seen as an endogenous entity the technology does not remain the same during the implementation process – something happens to the technology when it is applied by the users. This qualitative change is driven through numerous mutual intercourses between the technology and the users and is described as translations. The two different theoretical assumptions have methodological consequences, which will lead to different empirical activities and hence different conclusions. In the following this will be demonstrated in a hypothetical case of an EHR evaluation given by Tryggestad and Borum [18].

A Hypothetical Case

Imagine two different, but comparable, departments of internal medicine which have both implemented an electronic health record system. Both departments are evaluated by observation, but two entirely different results are obtained. In the first department (M1) medication errors were reduced by 20%. The length of stay was reduced by 10% for the patients whose records were computerized compared to a control group of patients in the same department whose information was kept in the usual paper record. In the other department (M2), no changes were found as a consequence of the EHR. The patients treated at the two departments were highly comparable, but there was a significant difference in the amount of human resources allocated. M2 had used 15% more staff time per treated patient than in M1. In M1 they only saw a marginal increase in the resources allocated using the EHR. The question for the two different organizational theories is what they can conclude.

Conclusions from the Use of the Diffusion Model

From the linear diffusion model the conclusion would be that the differences in the results could be explained by organizational issues: structure, communication or staff competence. The conclusion builds logically on the assumption that the technology implemented in the two departments are the same. Given this perspective the organizational issues will become the principal focus of investigation. The differences in meaning and significance of the technology to the two different departments will at the same time vanish, because the technology is regarded as a stable and continual entity for all users.

The very good results from M1 have shown that the EHR system has great potential and only marginal extra manpower is needed if the system is used correctly. The EHR system should be implemented in other departments where organizational issues such as competence building, appropriate training, workflow analysis, etc. should be given stringent attention.

Conclusions from the Use of the Translation Model

From a translation perspective the task would be to investigate how the implementation proceeded and find explanations to why it took place the way it did. The implementation process is reconstructed by means of interviews, document survey, etc.

In the hypothetical case it could be noted that the two departments had different relations to the vendor company. Over a number of years, the physicians at M1 have been invited to participate in meetings and conferences, and an earlier version of the EHR system was used in M1 as a pilot study. When the final version of the EHR system was implemented a couple of years later in both M1 and M2, technical advisers from the vendor company were also particularly active in M1 this time giving basic introductions to computers, closely following up on work planning, etc. The vendor company had limited resources for this implementation process, and the dominating part of the consultancy was allocated to M1 where they already had a good network.

Including the history of the system implementation provides further explanation on the observed differences between department M1 and M2. The EHR system is not the same technology to M1 as it is to M2. Through the relationship and interaction M1 has had with the vendor, they have created a unique technology. To M2 the interaction with the EHR system is a different story. M2 was not involved in testing of the earlier version of the EHR system, and the vendor is primarily engaged through M1 – consequently the EHR system does not have the same technology history in the two departments. The historical point is revealed because the translation model tries to uncover the technological history behind the sequence of events.

Given the analysis from a translation perspective, the final results of the evaluation would be slightly different. Results from the evaluation of the two departments are not unambiguous. Likewise, that would be a fact for an analysis of the factual cost of the implementation as substantial resources over time have been provided by the vendor company. The results can be explained by
the investigation of the technological history, which showed a significantly different sequence of events behind the implementation. Consequently, the EHR should not be recommended as the standard best solution, but should be evaluated in terms of the capacity of the departments that will be next in line for EHR implementation.

From the hypothetical case it is seen that two significantly different conclusions can be reached according to how the relation between organization and technology is perceived. From the diffusion model where the technology is seen as exogenous to the organization it is recommended to continue the implementation of the EHR system, whereas the translation model which recognizes the technology as endogenous to the organization includes the history of the technology to conclude that the EHR system is not necessarily the best. It should be further evaluated in terms of capacity parameters of the specific department.

Key Steps in an Evaluation Study

As mentioned it is impossible to present a general recipe for evaluation projects because EHR systems contain a number of different components, each taking care of a number of different functions in complex organizations. It is, however, possible to describe the design of evaluation projects as following several steps [19]. The steps are described as a number of decisions and agreements that have to be made, and as the outcome range of these decisions and agreements are wide, it is impossible to describe all the consequences of the outcome in all the steps. The first four steps are crucial because the outcome of these steps have consequences for the selection of adequate methods. The first step is to agree on why to evaluate. This step is crucial because in the decision on why to evaluate it is also determined what to do with the end result of the evaluation. If it is decided to evaluate an EHR system by comparing the situation before and after implementation, and the outcome of the evaluation indicates that the situation before the EHR was implemented is preferable, will the organization then be willing to close down the system again? In this first step of the evaluation it is essential to decide who is going to act on the eventual result of the evaluation and to agree to what extent they will act on the results.

The second step is to decide on what to evaluate – what is the primary object of the evaluation? The two most salient domains are functionalities or technical design. From the technical domain it could for instance be evaluated whether the EHR system is compatible with other systems, if it is easily maintained, does it have an acceptable response time, etc. From the functionality domain it could be evaluated whether the EHR system has the requested features, are they easy to use, or do they support the clinical tasks adequately?

The third step is to decide when to evaluate – at what stage of the life cycle will the system be evaluated. Evaluations in the design or development phase can often be performed in laboratories where it is possible to conduct a controlled environment. If the system is implemented or in the implementation phase it is usually more complex to evaluate as the system is a part of daily work routines in a busy clinic.

The fourth step is to decide on which stakeholder viewpoints to pursue. When evaluations are made to form the basis for decisions there will most often – if not always – be different viewpoints between the stakeholders. These differences can evolve to serious conflicts in the organizations if not dealt with appropriately. They should therefore be considered prior to the evaluation process, as the decision made will influence the design of the evaluation study.

The fifth step is the most difficult one – to decide on the methodology. It is essential to deal with the methodological themes that were presented above as opposing issues: a. Objectivist versus subjectivist approach b. Qualitative versus quantitative methods c. Reductionist versus holistic approach d. Summative versus formative evaluation.

There are clearly a number of interdependencies among the issues mentioned, and it is not possible to uncover them all and foresee all consequences. The best thing to do is to include a few iterations on the first five steps mentioned to reveal hidden contradictions and ensure interoperability. When it has been decided what, when and how to evaluate the sixth step is to collect the data, analyze them and report the results. How the data are to be collected and analyzed are mainly decisions that have been made under the methodological considerations, but the analysis can reveal aspects which are difficult to account for and explain. Furthermore it may be impossible to make valid conclusions due to contradictory or unclear results, and sometimes the health care staff will have concerns which cannot be shown in the collected data. Finally, EHR systems constantly change – not only are new versions of software regularly issued, but changes in work procedures occur continuously and training sessions increase the capacity of the health care workforce.

The evaluation results should be presented in a clear and concise manner; to ensure that the results are communicated to the different stakeholders, different reports must be prepared to reflect their perspective and language. The seventh and last step is to assess the results of the evaluation and decide on actions to be taken. This is an important part of the evaluation, but very often overseen. There is a clear reference to the first step where it was decided why to evaluate. The reason for evaluating should be stated clearly and in this reason there is also an argument for how to react to the results. Negative results are often neglected because it can be difficult to face the consequences, but evaluations should be regarded as part of a learning process and a source of new knowledge to improve EHR systems and the way they are implemented and used.
Conclusion

Evaluations can be quite complex to carry out, and despite good guidance in a number of fine textbooks and numerous references in the scientific literature there is no simple evaluation recipe. This article presents an overview of the various considerations one has to reflect on in the design of an EHR evaluation. A number of issues are presented and seven key steps are listed to guide the design of evaluation projects. At the same time it is also mentioned that the issues presented are not completely exhausted and the seven key steps might have to include iterative loops because of interdependence between some of the steps.

It is, however, important to consider all the mentioned aspects carefully in the attempt to meet the growing demands for evaluations of EHR systems. But we can only avoid being choked to death by the porridge of evaluation demands by making systematic and explicit choices and decisions.

References


Correspondence to: Christian Nøhr, M.Sc., Ph.D.
A/Professor
Department of Development and Planning
Aalborg University
Fibigertsræde 13
9220 Aalborg
Denmark
E-mail: cn@plan.aau.dk