HANDOUT: MEDINFO 2004

Workshop: "Good Evaluation of Health informatics Applications"

Jytte Brender, Aalborg University, Denmark

Abstract: This paper is an excerpt of a handbook of methods for assessment of IT-based solutions within healthcare, and which aims at facilitating evaluation specifically from the users' perspective. The scope is to smooth the progress of healthcare staff (or their representatives) carrying out assessment studies fully or partly themselves and at a satisfactory level of quality of the outcome. A series of recommendations on how to proceed with and accomplish an assessment study are briefly discussed, followed by tables listing a number of candidate assessment methods matching different information needs and distributed on the various phases of the development or implementation and operation of an IT-based solution.

1. INTRODUCTION

There is a big difference in what the developers and the users shall or must evaluate for a Health Informatics & Telecommunication System (in short referred to as 'IT-system'). This paper aims at facilitating evaluation specifically from the users' perspective. It is an excerpt of a handbook of methods for technology assessment of IT-based solutions in healthcare, achieved by the MUP-IT Project[1], [1]. The primary scope is to support the user in accomplishing an assessment study without too many scratches, and then to illustrate options for finding appropriate tools within the literature. Thus, the target readers of this paper constitute healthcare professionals of IT-systems within healthcare including their colleagues, consultants or other representatives. However, an evaluation method is not a thing that one pulls out of the sleeve, a drawer or even books and uses without reflection and meticulous care, and therefore, a list of desirable capabilities and personal qualifications are listed.

The handbook summarises current user-oriented assessment methods, for which the practicability suffice for a user organisation to carry out (parts of) the assessment itself with a satisfactory level of quality of information in response to a specific information need. Methods

[1)] The handbook of methods is essentially a synthesis of work carried out in R&D projects for more than a decade: The basis was formed during a number of Health Informatics Projects within the EU Commission's 3rd to 5th Framework Programme (primarily the projects KAVAS (A1021), KAVAS-2 (A2019), OpenLabs (A2028), ISAR (A2052) and CANTOR (HC 4003), as well as the concerted actions COMAG-BME and ATIM), followed by my PhD study financed by a grant from the Danish Technical Research Council. The final synthesis was enabled by a grant for the MUP-IT project by the Danish Center for Evaluation and Medical Technology Assessment.

requiring a special, professional background are excluded, such as statistical and econometric methods, theories and techniques.

1.1 What is evaluation?

'Evaluation' is - formally defined - "acts related to measurement or exploration of a system's properties". In short and a bit simplified 'system' means "a way of doing things"[2]. Evaluation may be accomplished during planning, during development, or during the operation and maintenance of an IT-system. When put in its logical extreme, 'evaluation' simply means to put numbers on some properties of the system, and consequently, 'evaluation has no sense as a self-contained and independent activity: The purpose of evaluation is to provide the basis for a decision about the investigated IT-system in some decision-making context, and that decision-making context is also the context for the evaluation:

> *"Evaluation can be defined as the act of measuring or exploring properties of a health information system (in planning, development, implementation, or operation), the result of which informs a decision to be made concerning that system in a specific context"*
>
> (citation from [2])

When evaluation is used in the context of controlling whether an IT-system - for instance at delivery - fulfils a prior agreement, then the activity is called 'verification'. Similarly, the concept of 'validation' is used, when the decision-making context is concerned with whether or not the system suffices to fulfil its purpose. The concept 'assessment' is used as a collective term, when it is unnecessary to distinguish between the different types of purpose with an evaluation.

When addressing the issue of assessing an electronic healthcare record or another IT-system within healthcare, the object for the assessment activity is usually the entire organisational solution and not only the technical artefact[3]. Thus, evaluation of an IT-based solution is concerned with not only the IT-system, but also its interaction with its organisational environment and its mode of operation within the organisation. This includes for instance actors (physicians, nurses, and other kinds of healthcare staff, as well as patients), work processes and structured activities, as well as external stake holders, and last but not least a mandate and a series of internal and external conditions for the organisation's operation. The assessment activity must act based on this wholeness, but of

"The recognition that evaluation has covert as well as overt functions necessitates its treatment as a social activity rather than the purely formal procedure it was previously held to be. .... we view evaluation not as the application of a set of tools and techniques, but as a process to be understood. By which we mean an understanding of the functions and nature of evaluation as well as its limitations and problems."

4

[2] "all the components, attributes and relationships needed to accomplish an objective" [3].

[3] The paper distinguishes between an 'IT-system' and an 'IT-based solution'. The term 'IT-system' denotes the technical construct of the whole solution (hardware, software, including basic software, and communication network), while 'IT-based solution' refers to the IT-system PLUS its surrounding organisation with its mission, conditions, structure, work processes, etc.

course limited to what is relevant in the specific decision-making context.

The above short introduction to concept of ‘evaluation’ indicates - as the citation in the text box - that to accomplish an evaluation requires more than picking a standard method and using it. The important is to understand the process within which the future result is going to be used, as well as the premises for obtaining a good result - before one at all chooses a method. Applying a given method in reality probably is the easiest yet laborious part of evaluation.

1.2 Required skills to be able to evaluate an IT-system appropriately

In the below the personal and professional qualifications needed for accomplishing assessment studies are briefly discussed. In the end, this of course depends of the specific information need and the complexity of the overall system.

- It is necessary to be able to disregard the political interests of one’s own profession, while viewing the situation in a larger perspective. There is a need for the ability to obtain an overview and thereby a need for making abstractions and reflections about issues within their larger context, before, during and after an assessment study. Often, activities may be carried out in more than one way, and it is necessary to be able to judge which one is the best given the conditions
- It is necessary to have skills for dealing with methods. One must be able to and dare forming one’s own opinion and carry it out, for instance in terms of variations of known methods. It is also important that one is capable of capturing, handling and interpreting deviations within the actual investigation and among the factual observations. Such deviations are more or less inevitable in practice
- There is a need for stringency and persistence, as assessment studies usually are large and imply a hell lot of data, the analysis of which must be methodically, cautiously and exhaustively carried into effect
- It is necessary with insight into IT-based solutions at a level enabling one to see through the technology, its conditions and events, and so that one dares setting boundaries (without being stubborn) towards professionals with formal credentials as better qualified in a certain respect. It is necessary that one is able to see through the implications of the interaction between the IT-systems and their organisational environment, as well as the organisation’s (and that of its individuals’) conditions, interactions and events
- There is a need for skills and experience as project manager. An assessment study is not something that one does single-handedly or as a desk test. Therefore, an ability is required to coordinate and delegate, to perform critical problem solving or resolution and to predict the consequences of one’s decisions and initiatives
- There is a need for the courage and the desire it takes to make oneself acquainted with new material and to search information at places where one normally does not move; in literature, on the Internet, at professionals in other disciplines, …
- It is necessary to be able to remain constructively critical, to the limit of suspicion, towards what is stated verbally or in writing, including one’s own approach and results
- There is a need for a thorough insight into the healthcare sector, including the conditions that such organisations work under

In short, it is important with the capability of a responsible and self-dependent thinking, with an overview, insight and an ability to reflect, including the ability of being self-critical.

2. TYPES OF ASSESSMENT OF IT-BASED SOLUTIONS

What is it that makes the difference between different assessment investigations? And why is it difficult to identify articles within the literature that one can benefit from? It is the degree of complexity of the domain of assessment as shall be briefly illustrated in this chapter.

2.1 Evaluation in a holistic perspective

There are many different and roughly independent ways to categorise evaluation, called dimensions. For instance there are two main focuses at user assessment:

- Concerning the technical construct (i.e. whether the construction of the IT-system is done satisfactorily)
- Concerning the functional mode of operation (i.e. whether the IT-system has the right functions and execute these in an appropriate way)

Orthogonal to this dimension, there are two main ways of assessing:

- Explorative, i.e. for clarification purposes (within the social sector this is called 'scope-free evaluation'). The aim at this type of assessment is to obtain a rich picture of what takes place in an organisation, how or why; in this situation there is typically no frame of reference. Psychological and sociological investigations are often of this type
- Verifying, i.e. of a confirmatory nature. This includes validation of for instance objectives fulfilment or attainment. At this type of assessment investigation one always uses a frame of reference (such as a contract or a specification) for comparison. Natural science investigations are often of this type, that is, an investigation is founded on a specific understanding (a model and/or a hypothesis), which the researcher attempts to reject

On top of the above, assessment may be either prognostic or diagnostic (including screening) or monitoring, while they cannot take a treatment nature, because of their nature as provider of a decision-making basis.

The spectrum of assessments adds further to this complexity and ranges from:

- Quantitative to qualitative. This is concerned with whether a measuring scale is available for putting the results into a metric context, or, as the other extreme, whether the result is purely descriptive
- Objective and subjective: Objectivity is concerned with rational measurement of physical properties (like response time in a dialogue between a user and the IT-system), while subjectivity is concerned with aspects that are emotionally based (like user satisfaction

and quality of job). It is not always feasible to separate these aspects from those within the next bullet, se for instance [5], however, this separation is preferred by the author.

- Reductionism and holism: Reductionism is an expression for the perspective[4] that everything consists of parts, which may be taken out of their context and measured or otherwise investigated individually, and that one afterwards may conclude about the wholeness on the basis of the outcome from individualised measurements on its component parts. The holistic perspective is based on the understanding that the wholeness comprise not only its component elements, but also other aspects that in a mutual interaction contribute to the creation of the wholeness. Synergy among members of a team is an example of this add-on element. The holistic perspective implies that one as an investigator cannot separate the human factors from the understanding of an organisation and its IT-based solution; see for instance myth no 1 in [8].

It would be inaccurate for the above-mentioned aspects to say that one extreme is better than the other; they answer different questions when applied. Further, if it is not feasible to accomplish a quantitative and objective investigation, one may as well get valuable information from a subjective and qualitative investigation. Still, one should not answer a quantitative question by a qualitative method. Similarly, one should not unconditionally reject the reductionistic approach that has been dominant within natural science research for centuries, because this may sometimes be the only way forward with an investigation. The important is that one realises the guiding perspective behind ones statements and choices as well as the pitfalls and perils each method suffer from.

Still another dimension in the understanding of ways to assess IT-based solutions is the distinction between constructive and summative assessment:

- Summative assessment has the purpose of contributing with a kind of status in a decision-making context. Typical examples of *summative assessment* are 1) evaluation of objectives fulfilment, or 2) the kind of assessment carried out in a contractual relation, when an IT-system is delivered and one wants to check, whether the installed system operates in compliance with the contract
- 'Constructive assessment' or 'Formative assessment' is the kind of assessment activities that aiming at giving directions for the decision-making concerning subsequent development or implementation tasks: most IT-project constitutes a compromise between

---

[4)] The concept of 'perspective' stands for hidden aspects and assumptions deeply buried in the design and application of methods, see for instance [6, 7]. In a generalised version, the perspective is the implicit assumptions on (cause-effect relations within) the object of study. So, the perspective is synonymous with "that aggregation of (conscious or unconscious, epistemological) assumptions of how things relate in combination with imprinted attitudes guiding our decision making e.g. in a problem solving situation". In an organisational context this corresponds to "assumptions on the nature of the working processes and how people interact in an organisational setting", i.e. the implicit and tacit models behind the principles that form the basis of our development and management theories. Thereby, these models are at the bottom of how we deal with things. Few method designers are aware that our cultural background (professional, religious, and national) alone maintains a series of tacit assumptions affecting our way of doing things; see for instance [6] page 6ff.

an ideal solution and something realisable, controlled by local concerns, considerations and limiting factors. Further, irrespective of whether it is an off-the-shelf system or not, most IT-projects are both large and unpredictable, among others because of organisational conditions and changes. Such dynamically changing conditions for assessment imply huge demands on handling evaluation methods for the establishment of an ongoing decision-making basis

At summative assessment of the functional mode of operation it is usually taken for granted that when the users sit down by the IT-system in order to evaluate it, it is (reasonably) free of programming bugs. However, this petition is not feasible at constructive assessment, where the constructive assessment is used precisely to guide the future direction of the development. Instead one has to handle errors as a necessary evil.

2.2 Evaluation viewed in the light of the IT-system's life cycle

The life cycle for an IT-based solution is the time interval ranging from conceiving the idea of an IT-based solution to meet a given intent and until the final solution is abandoned. The solution may concern a complete IT-system, but may also just concern a specific change in an existing IT-system.

An IT-system's life cycle consists of a series of phases, where according to the international standardisation organisation, ISO, the concept of 'phase' means "a defined segment of work" [9]. This definition suggests that the distribution in time for activities in a given phase may well be accomplished in parallel with activities of other phases. At the same time the definition signifies that there may be other equally useful ways of dividing the time axis into phases for the IT-systems life. The argument is that it depends on the global approach and the methods chosen to accomplish the major tasks, such as the method of analysis that lead to a requirement specification: A spiral approach has a different set of phases than a waterfall model for systems development, or the same concepts are interpreted differently.

The present life cycle model merely has the purpose of providing a framework for linking evaluation methods to typical tasks and activities in a systems development or implementation project. For this purpose and irrespective of the system development methodology, we will use four superior phases for the life cycle of an IT-based solution, as outlined in the following.

The ***Explorative Phase*** is concerned with the strategic aspects related to the global task, addressing issues like objective, intentions with the solution, basic principles for the solution model as well as feasibility of realisation (technically and organisationally). This phase typically ends with a requirement specification and a contract or another kind of specification of what the users want to achieve.

The ***Technical Implementation Phase*** comprises the development and/or adaptation of the IT-system, followed by its installation within the organisation, also in terms of changed working procedures. User assessment in this phase typically constitutes constructive assess-

ment activities, concluded with a summative verification addressing whether the IT-system/IT-based solution is good enough to be put into daily operation, and whether the contractual agreement is fulfilled.

The ***Adaptation Phase*** (previously in [10] called the Maintenance Phase, see Figure 1) includes the period immediately after the IT-system fully or partly has been taken into daily operation. It is a phase where extensive adaptations of the IT-system, its set-up and/or the working procedures take place, until a stable situation is achieved. It is assumed that all significant errors (bugs) and omissions found in the previous phase have been corrected prior to the daily operation. This phase therefore, is concerned with corrections, adaptations and refinement of all types of errors, omissions and shortcomings identified after the system has been taken into operation, and which are necessary for the organisation to become effective (again).

The ***Evolution Phase***, at which the worst troubles in the IT-systems and its surrounding work procedures have been overcome, and the operation has entered into a state of (reasonable) stability. During the Evolution Phase a number of new development or change activities will usually be initiated, each of which will have a life cycle corresponding to the present life cycle model. Also measures of the effect or impact of the IT-system or solution may take place in this phase. It ends when the IT-system/solution is replaced with another system or solution.

The figure below shows the dynamic aspect in assessment, related to the present life cycle model. The impact of constructive assessment is illustrated in terms of arrows pointing forwards and backwards to other phases, while the summative assessment isn't visible in terms of arrows as it is totally contained within a given phase. For constructive assessment the figure therefore shows the entire complexity and illustrates why assessment may be a task for dedicated professionals with experience. The arrows comprising feedback loops (pointing upwards in the downwards directionality of overall progress) illustrate that some decisions implied by the outcome of an assessment may be radical and may even imply changes in previous decisions.

Even if the summative assessment from an external view seems rather simple and static, the constantly changing situation also complicates such assessment tasks. For instance, if one has made a baseline investigation to be used as the frame of reference for a planned later assessment study, one may still risk that the foundation for the study changes, implying that the frame of reference becomes more or less invalid or inaccurate: Normally the implementation of an IT-based solution takes several years from defining ones' requirements, getting the budget allocated, finding the right vendor, and until the chosen system is taken into daily operation, - irrespective of whether it is an off-the-shelf system one buys. Within this period a number of aspects may change drastically: departments may fuse or restructure; new technology may appear for one or more of the department's normal activities; the department may get more, new or changed tasks delegated, or it may be the victim of general budget cuts; etc. All of this may change the conditions for an assessment investigation. In particular the frame of reference is vulnerable in connection with certain summative assessment investigations.

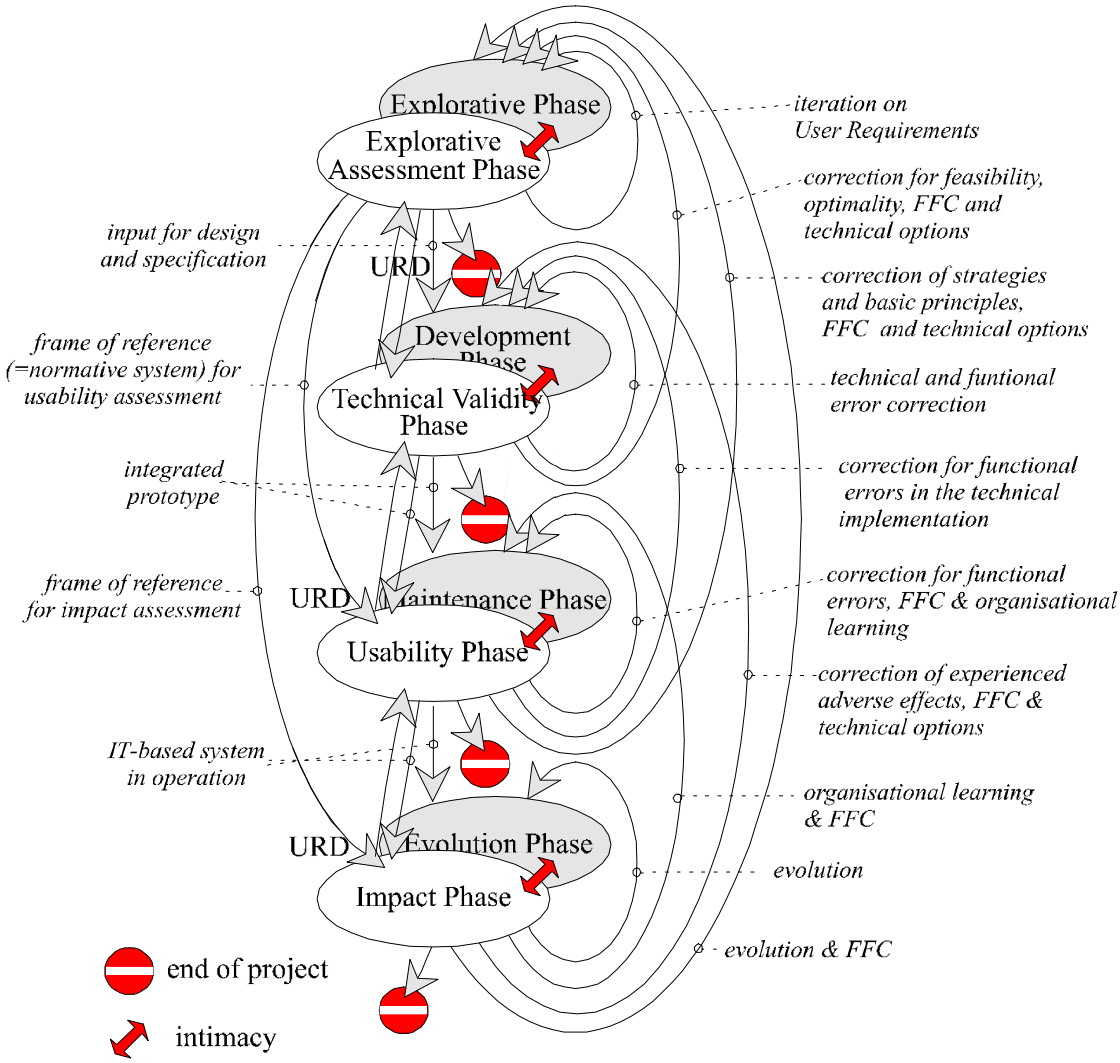


Figure 1: The Dynamic Assessment Methodology, completed with descriptions of feed-forward loops (providing frames of reference and preventive measures) and feed-back-loops (initiating corrective activities) and indicating the contents of this information flow. The hatched ellipses illustrate the technical activities, whereas the white ellipses illustrate the corresponding constructive assessment activities in a four-phased structure. The thick arrows between the technical development and assessment activities indicate unspecified interaction (co-ordination and collaboration). (URD = User Requirements Document; FFC = Four Founding Capacities, i.e. characteristics regarding the capability and capacity of accommo-dating changes), from [10].

It is all of the above that the many arrows within Figure 1 illustrate.

The figure also illustrates that it may be risky stringently to stick to a plan just because it was once decided upon. The reason is - again - that the conditions for assessment may have changed considerably: The resources allocated may have gone in reality or relatively. For instance, the latter is the case if things turn out more expensive than anticipated at the planning. It may be that one is capable of catching up parts of it, or that the residual study

may be rationalised to achieve a higher effectiveness. However, at least be aware that it might be more beneficial to move a deadline rather than sticking to it. Another option might be to reduce the level of ambition in case the project turns out being more wide-ranging or costly than anticipated.

3. HOW DO YOU DO? - ABOUT CHOOSING OR CONSTRUCTING METHODS FOR EVALUATION

First,

"(1) There is no *one* way to do evaluation.
(2) There is no generic logical structure that will ensure a unique right method of choice.
(3) Evaluation ultimately becomes judgment and will remain so, so long as there is no ultimate criterion for monotonic ordering of priorities; and
(4) …."

(Weckwerth, cited in [11]),

The author unconditionally agree to point (1) and (2), see the discussions in [7, 10]: It is not yet feasible to make a cookbook for assessment, and …

> *"we view evaluation not as the application of a set of tools and techniques, but as a process to be understood. By which we mean an understanding of the functions and nature of evaluation as well as its limitations and problems."*
>
> 4

... is still a valid statement. Thereby, also (3) is valid.

A series of factors determine the success of an assessment project, and this is what this chapter will outline.

3.1 Where in the life cycle is the IT-project?

The very first issue one has to consider is where the assessment study is positioned in terms of the life cycle of the IT-system, i.e. its development or implementation stage. Based on this information, one may by means of the overview tables listing methods identify some candidate methods (see Chapter 4), or one may get inspired for one's own further progress.

3.2 What is the information need?

There is a clear correlation between the actual information need and applicable methods[5]. Therefore, next step is to identify and delimit the strategic objective of the investigation, dis-

[5)] Examples of this relationship are described in [12, 13] – incompletely however, yet perfect to bring an understanding as well as inspiration for the further progress in planning.

cussing what it is that you really want to know. The essential in this respect is that there is a complete match between the purpose, the approach and the actual use of the result.

Consequently, if you don't know the purpose of evaluating, don't evaluate!

The next question is whether the outcome of the investigation is intended to ….

- Establish the foundation for:
  - The ongoing work in the development project?
  - An administrative or a political decision-making?
  - A healthcare professional or IT-professional decision-making?
- Establish arguments in a sales promotion of the IT-system? A good assessment study can always be used in a marketing context, provided that the system is good (enough) and one dares being objective
- Establish the basis for a scientific publication? Be careful when starting with one of the other scopes and subsequent trying to write a scientific article based on the outcome of the investigation
- Find out something? That is, exploration of the IT-based solution with the aim of investigating what changes are happening, have happened or are likely to happen within the organisation. Very often this scope has a research purpose

The first two scopes listed do not normally require the same accuracy and stringency as for scientific research. The risk from aiming in parallel at more than one of the above scopes is either that one falls between two stools, or that one cracks a nut with a sledgehammer. However, with caution it may be feasible.

Write down the information need and the purpose of the assessment as the starting point of the subsequent planning. It should be published in a way that allows everyone access and the ability of referencing to the information. An assessment study often requires extra efforts from the staff, and therefore, it is important that they understand the task and are loyal towards it. One gets the best results, if there is transparency, and everyone works based on motivation towards the same goal as the one delegating and authorising the task. Vigorous discussions may arise in connection with such clarification, as the issue is about the reasoning behind carrying out the investigation and hence for taking on the extra burden, i.e. the motivation.

It is also important that one realises whether one's object of study is the IT-based wholeness within the organisation, or any part hereof, including the technical construct?

3.3 About establishing a methodology

While methods constitute specific prescriptions for how to accomplish a certain task, a methodology is the superior wholeness bringing different methods to work together, also enabling a method distributed over several phases to function in a practical wholeness. In the light of major IT-projects, selecting a methodology will lead to the global plan for division of the

project into phases as well as the overall approach for each of the sub-tasks, the plan for user participation, for method selection, etc., or the overall approach for a before-and-after investigation of the system's effect on specific parameters of the organisation.

The selection of a methodology is usually implicitly described within the original description of a method. Establishment of a methodology comprises taking a well organised step from knowing the information need, over the conditions for carrying out the study (including policy for the accomplishment and the solution), to value norms, via consequences and assumptions that one must contemplate to an action plan. The methodology hereafter defines certain limits for choice of methods. If you have a simple, small and very specific information need, for which you may apply a specific, well-documented method having a detailed prescription, the methodology may not need to be explicit. Otherwise, it may be wise to think in terms of the wholeness.

The point of time within the life cycle is a significant factor for the selection of one's methodology: if one has entered the Adaptation Phase before even considering measuring something, then forget making a before-and-after investigation for identifying the effects of the IT-system on the organisation and its service level, unless there exists a baseline investigation of some sort, which may serve as the frame of reference for interpretation.

The nature of an assessment study, i.e. the distinction between constructive and summative assessment, is the first and crucial issue at the establishment of a methodology. Except form development projects, most assessment studies will be of a summative nature. These are usually larger, but also easier, because there will be enough of time for a careful planning. At constructive assessment, one has constantly to be at the leading edge, also being prepared for adapting oneself to the specific needs, while obeying the timelines and the budget conditions as well as the responsibility and authority conditions within the project.

Anticipated areas of effect - in combination with whether it is constructive or summative assessment - influence the choice of methodology and methods. Where the introduced technology inevitably and significantly affects responsibility and authority aspects of an organisation, i.e. changes the structure, relations, coordination or collaboration transversely within an organisation, it is obvious to address management issues. Oppositely, if a purely 'mechanical' effect were anticipated for the work procedures, methods dealing with work procedures would normally be in focus. If one aims at drawing a lesson from the investigation, a third type of methods would be relevant. And so on.

3.4 About choosing a method

Be critical at the selection of method(s):

> … Pick methods that comply with the objectives of the project (purpose of the assessment, including the need for accuracy and precision). Make a distinction between a practical and down to earth, internal purpose and a more formal, external-oriented purpose, including those initiated as a scientific investigation

... Be aware that one should NOT choose to carry out a qualitative investigation, only because there is not enough energy or budget to carry out the quantitative investigation, which the information need points at

… Choose methods that meet the terms of those operating them. It makes little sense to pick methods that technically or professionally are too difficult for those intended to carry out the investigation

… Choose methods for which the perspective are in harmony with your own perspectives, - otherwise it may be difficult to interpret the results; see the term 'perspective' in a previous footnote. Everyone has a series of assumptions, which we are indoctrinated with during our education and professional environment, and which unfortunately are tacit. The best to do, therefore, is carefully getting acquainted with the methods, before one makes the final choice

… Preferably choose methods that are validated, and which have documented their advantages and drawbacks within the literature. However, note that "being validated" does not necessarily mean the same as being valid for your purpose. One has to subject the original literature to a minute inspection and reflection, as only the best articles thoroughly analyse and criticise their method adequately

… Carefully check the list of perils and pitfalls in [14], so that no bias is overlooked - provided that a high level of accuracy and precision is needed

Then note that methods usually applied for analysis and design work in connection with system development some times may be valuable for assessment tasks also, especially for qualitative and explorative investigations. This is the case for instance, where there is a need for finding out how things are actually working as compared with the anticipated mode of operation, for instance in terms of an investigation of effect or objectives realisation.

One may easily apply a method wrongly, and therefore it is important that one prior to use has insight into the method, i.e. comprehending it at an abstract level. Don't just jump into using a method from its prescription. It is the understanding that facilitates a meta-level (birds' eye view), enabling one to juggle with and within the frame of a method as well as to interpret the results correctly and completely, compared to the conditions of the study and the data gathered. A method is not something that one picks out of a drawer to use. Exactly the same way that a cook (the chef in a restaurant) usually gets a better and quicker result out of a given prescription; the chef knows which qualities of the raw material and the preparation that makes the difference, and (s)he may take shortcuts, where normal people would fail.

3.5 About choosing metrics and measures

There is a close connection between metrics[6], measures and choice of methods, and this has to be respected. Every method has its focus area and typically leads to certain forms of information, both of which requires special expertise to change.

[6)] The concept of *'metric'* – for assessment methods – concerns specific measuring techniques or tools, but also often just a calculation formula or an instrument for measuring something specific. Consequently, it is the metrics that provide the *'measures'* with specific and reliable values. An

Select measures carefully:

… Focussing on visible measures bears a risk of implying a changed behaviour or performance by the employees. It is a well-known phenomenon already from the early era of assembly lines and mass production that people change, when they know that they are observed, - the Hawthorne Effect

… Fixation on specific measures may imply that these get special attention by the staff, causing a focus on them during the daily practise at the expense of the wholeness. This corresponds to a long-distance runner training and training his legs while forgetting other muscles; - the result of which may be an imbalance

… One may often find invisible measures. For instance, it may be as simple as retrospectively monitoring the consumption of coffee or the length of the coffee breaks before and after the introduction of an IT-system as a measure of its effectiveness. Or, if one wants to look at the communication effectiveness one may register the number of phone calls between specific numbers. This kind of information may be elicited without the users noticing

3.6 About execution of the method

Keep an eye on what happens within the organisation while you are assessing. That way you will know whether something (unforeseen) has to be taken into account during the analysis and interpretation of the results:

… An organisation changes: as soon as its staff get access to other or new tools and they are confident with these, they will invent short cuts. However, not all of these shortcuts will be compliant with the prescribed procedures, and some may even be contradictory to the interests of the organisation

… The indicators change: what one believed to know about certain variables used at the investigation may suddenly turn out having unknown limitations. This may be caused by the users' dynamically changing way of using the system, or previously unknown aspects of the IT-system. Or the work procedures may cause the indicators to change: It is not normal to 'work according to the rules', but the introduction of an IT-system may temporarily force this

… One's scope may change while accomplishing the assessment: an answer to one question may lead to ten new questions. Digging into a topic makes one wiser and that alone may change one's information need

… The IT-system with guarantee will change while being assessed, at least for newly developed IT-systems with correction of bugs, or if the assessment lasts for a period of months or more

To some extent the method can/shall keep pace with the above-mentioned changes. It is necessary now and then to introduce modifications, as applicable. The art of assessment means

---

example of a metric of relevance for a given Electronic Healthcare Record (EHR) project is the formulae for calculating a number of different measures addressing the effect of an EHR on the quality of a clinical departments service.

having a sufficient bird's eye perspective to be able to know and understand when and how it is feasible to change or correspondingly accept that things change by themselves, without ruining the outcome of one's investigation. The art is also to observe, note and deal appropriately with such changes. Be honest about them in a publication context, as some of them may affect the interpretation of one's results and conclusion, and thereby also influence others' benefit from using the same method.

3.7 About interpretation of results

Be critical with what you have your name associated with:

… Be fair and objective, be open towards unwanted and unexpected results, and accept when relevant their implied consequences. It may sometimes be hard to face unwanted or unexpected results, but for constructive assessment this is essential

… Be self-critical and show that you are aware of the weaknesses of both the method and the interpretation. This is worth its weight in gold for others considering applying the same method or IT-solution, as it shows that you are so much in control of the investigation that the outcome can be trusted

… Be objective while presenting the results, i.e. showing both the good and bad news at a relevant level of detail and in the right mutual proportion. Exclusion of significant information in the synthesis of a conclusion may strain belief or even come near to fraudulence

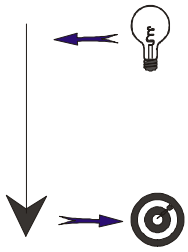
4. OVERVIEW OF EVALUATION METHODS

In principle all aspects are candidates for assessment in all phases of a system's development. In practice, some aspects are more prominent in some of the phases than in others. Within the life cycle, assessment may change in nature from being prognostic (during planning), to screening and diagnosing (prior to shifting to daily operation), to treating (at the handling of known error situations or shortcomings). Note, therefore, that even if a method is not listed under a specific phase, an information need may arise that requires inspiration from the methods listed under other phases.

Note that few of the references given include a discussion of their weaknesses, perils and pitfalls.

4.1 Overview of assessment methods: Explorative Phase

The methods included in this section are particularly relevant at assessment of issues raised during the establishment of a User Requirements Specification, like objectives, requirements and expectations.

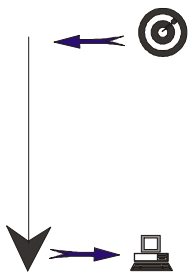


Method	Application area	References
Analysis of work procedures	Elucidation of how things are actually carried out within an organisation	(search for systems analysis methods)
Assessment of bids	Comparative assessment of a number of offers from one or more bidders/vendors	[1]
Balanced Scorecard	Ongoing optimisation of the outcome of a development project by balancing focus areas by means of a set of indicators for a set of strategic objectives	[15-16]
BIKVA	Critical, subjective assessment of an existing practice	[17-18]
Delphi	1) (Qualitative) assessment of an effect, for instance where the solution space otherwise is too big for handling 2) Exploration of development trends 3) Elucidation of a problem area, for instance prior to a strategic planning	[19-22]
Field study	Observation of and within an organisation to identify its practice and to elucidate mechanisms controlling change	[23-24]
Focus Group Interview	Is in principle used for the same purposes as other interview methods. In practice, the method is most relevant during the early analysis phase, for instance where attitudes or problems of social groups needs elucidation, or when a model solution is being established	[25-28]
Future workshop	Evaluation and analysis of an (existing) situation with respect to identification of focus areas for change, i.e. aiming at designing the future practice	[29]
Grounded Theory	Supportive analytical method for data acquisition methods that generate textual data, such as some open questionnaire methods and interviews (individual and group interviews)	[30-32]
Heuristic evaluation	Is used, when no other realisable possibilities exist, for instance when: ▫ The organisation does not have the necessary time or expertise ▫ There are no formalised methods ▫ There is not yet something tangible to assess	[33]
Interviews (non-standardised)	Is in particular suited for elucidation of individuals' opinions, attitudes and perception regarding phenomena and observations	[34-36]
KUBI	Optimisation of the outcome of a long-term development project, based on a set of user or customer/client defined value norms and objectives	[18]
Logical Framework Approach (LFA)	Situation analysis to support at the choice of focus for a development, but at the same time a simple technique for incorporation of risk handling within a project planning	[37]

Pardizipp	Preparation of scenarios of the future	[38]
Questionnaires (non-standardised)	Questionnaires are used to answer a wide range of questions, but for investigations requiring a high level of accuracy, its main application area is (qualitative) investigations of subjective aspects	[39-43]
Requirements assessment	Within the European culture the User Requirements Specification is the basis for purchasing an IT-based solution or engaging in a development project. Consequently, the User Requirements Specification is a highly significant legal document, which needs thorough assessment	[1]
Risk assessment	Identification and subsequent monitoring of risk factors, enabling the possibility of taking actions before an unfortunate situation develops	[44]
Social network analysis	Assessment of relations among elements within an organisation (such as individuals, professions, departments or other organisations), which influence the acceptance and use of an IT-based solution	[45]
Stakeholder analysis	Assessment of stakeholders' features and their inner dynamics, aiming at the identification of participants for the accomplishment of a given task, problem solving or project	[37, 46]
SWOT	Situation analysis: establishment of a holistic view of a situation or a model solution	[1]
Usability	Assessment of user friendliness in terms of ergonomic and cognitive aspects of the interaction (dialogue) between an IT-system and its users. In this phase the concern is a planning or purchase situation	[47-50]
Video recording	Monitoring & documentation as a means for analysis of what/how the work procedures respectively the users' activities are actually carried out, or for investigation of complex patterns of interaction	[51-52]
WHO: Framework for the assessment of strategies	Assessment of different (development) strategies either individually or as a comparative analysis	[1]

4.2 Overview of assessment methods: Technical Development Phase

The methods listed in this section are particularly suited for user activities during the development and installation of an IT-based solution and may be used to provide feedback to the technical development.

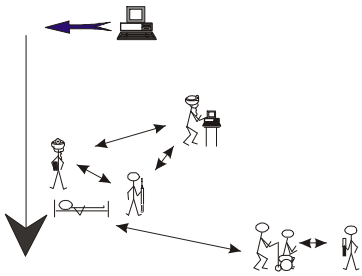


Assessment in this phase is typically carried out under experimental conditions, and not during real operation. The phase is usually completed with a technical verification, to make certain that all necessary functions and features are present and work properly in compliance with the established agreement.

Method	Application area	Refe-rences
Balanced Scorecard	Ongoing optimisation of the outcome of a development project by balancing focus areas by means of a set of indicators for a set of strategic objectives	[15-16]
Clinical/diagnostic performance	Measurement of diagnostic 'correctness' (for instance measures of accuracy and precision) for IT-based expert systems and decision support systems	[53-55]
Cognitive assessment	Assessment of cognitive aspects of the interaction between an IT-system and its users, for instance: ▫ Identification of where and why operational errors occur ▫ Identification of areas for a focused improvement of the user friendliness	[51-52, 56]
Cognitive Walkthrough	Assessment of user 'friendliness' on the basis of a systems design, from specifications, muck-ups or prototypes of the system, aiming at judging how well the system complies with the users' way of thinking, for instance: ▫ Identification of where and why operational errors occur ▫ Identification of causes behind problems with respect to the user friendliness and consequently identification of areas for improvement	[57-58]
Heuristic evaluation	Is used, when no other realisable possibilities exist, for instance when: ▫ The organisation does not have the necessary time or expertise ▫ There are no formalised methods ▫ There is not yet something tangible to assess	[33]
Risk assessment	Identification and subsequent monitoring of risk factors, enabling the possibility of taking actions before an unfortunate situation develops	[44]
SWOT	Situation analysis: establishment of a holistic view of a situation or a model solution	[1]
Technical verification	Verification that the agreed functions are present, work correct and in compliance with the agreement. This may take place for instance in connection with delivery of an IT-system or prior to daily operation, and at any subsequent change of the IT-system (releases, versions, and patches)	[1]
Usability	Assessment of user friendliness in terms of ergonomic and cognitive aspects of the interaction (dialogue) between an IT-system and its users	[47-50]

4.3 Overview of assessment methods: Adaptation Phase

Evaluation in this phase has the purpose of providing support for the modification or refinement of the IT-based solution, work procedures as well as functions implemented within the IT-system, to make the two of them work optimally as a whole during daily operation. This phase should be fairly short, provided that the implemented solution is well functioning from the beginning.



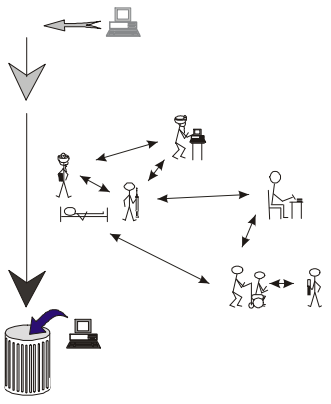
Now that assessment can take place in real operation, ergonomic, cognitive and functionality assessment will gain much more focus, as potential inadequacies or shortcomings in this respect will present themselves as operational errors, misuse and alike.

Method	**Application area**	**References**
Analysis of work procedures	Elucidation of how things are actually carried out, in comparison with the expected. This includes its use in relation to measures of the degree to which the IT-system is used as anticipated	(search for systems analysis methods)
BIKVA	Critical, subjective assessment of an existing practice	[17-18]
Clinical/diagnostic performance	Measurement of diagnostic 'correctness' (for instance measures of accuracy and precision) for IT-based expert systems and decision support systems.	[53-55]
Cognitive assessment	Assessment of cognitive aspects of the interaction between an IT-system and its users, for instance: ▫ Identification of where and why operational errors occur ▫ Identification of areas for a focused improvement of the user friendliness	[51-52, 56]
Cognitive Walkthrough	Assessment of user 'friendliness' on the basis of a systems design, from specifications, muck-ups or prototypes of the system, aiming at judging how well the system complies with the users' way of thinking, for instance: ▫ Identification of where and why operational errors occur ▫ Identification of causes behind problems with respect to the user friendliness and consequently identification of areas for improvement	[57-58]
Field study	Observation of and within an organisation to identify its practice and to elucidate mechanisms controlling change	[23-24]
Focus Group Interview	Is in principle used for the same purposes as other interview methods. In practice, the method is most relevant during the early analysis phase, for instance where attitudes or problems of social groups needs elucidation, or when a model solution is being established	[25-28]
Functionality assessment	1) Validation of objectives fulfilment (realisation of objectives), i.e. the degree of compliance between the desired effect and the materialised solution 2) Effect assessment (also called: impact assessment) 3) Identification of problems within the relation between work procedures and the IT system's functional solution The method will capture severe ergonomic and cognitive problems, but is not dedicated at capturing details of this type	[10]
Grounded Theory	Supportive analytical method for data acquisition methods that generate textual data, such as some open questionnaire methods and interviews (individual and group interviews)	[30-32]

Heuristic evaluation	Is used, when no other realisable possibilities exist, for instance when: ▫ The organisation does not have the necessary time or expertise ▫ There are no formalised methods ▫ There is not yet something tangible to assess	[33]
Interviews (non-standardised)	Is in particular suited for elucidation of individuals' opinions, attitudes and perception regarding phenomena and observations	[34-36]
Questionnaires (non-standardised)	Questionnaires are used to answer a wide range of questions, but for investigations requiring a high level of accuracy, its main application area is (qualitative) investigations of subjective aspects	[39-43]
RCT, Randomised controlled studies	Verification of efficacy, i.e. that the IT-system – under ideal conditions – makes a difference for the patient care. In particular used for studies of decision-support systems and expert systems	[55, 59]
Risk assessment	Identification and subsequent monitoring of risk factors, enabling the possibility of taking actions before an unfortunate situation develops	[44]
Social network analysis	Assessment of relations among elements within an organisation (such as individuals, professions, departments or other organisations), which influence the acceptance and use of an IT-based solution	[45]
SWOT	Situation analysis: establishment of a holistic view of a situation or a model solution	[1]
Technical verification	Verification that the agreed functions are present, work correct and in compliance with the agreement. This may take place for instance in connection with delivery of an IT-system or prior to daily operation, and at any subsequent change of the IT-system (releases, versions, and patches)	[1]
Usability	Assessment of user friendliness in terms of ergonomic and cognitive aspects of the interaction (dialogue) between an IT-system and its users	[47-50]
User acceptance and satisfaction	Assessment of users' opinion, attitudes, and perception of and IT-system at daily operation	[1]
Video recording	Monitoring & documentation as a means for analysis of what/how the work procedures respectively the users' activities are actually carried out, or for investigation of complex patterns of interaction	[51-52]

4.4 Overview of assessment methods: Evolution Phase

This phase is usually considered starting at the point in time where the entire IT-based solution has reached a state of sufficient stability with respect to bugs and corrections, and when evolutionary activities are started up. Consequently, the shift between this and the previous phase may be fluid.



Method	Application area	Refe-rences
Analysis of work procedures	Elucidation of how things are actually carried out, in comparison with the expected. This includes its use in relation to measures of effect	(search for systems analysis methods)
Balanced Scorecard	Ongoing optimisation of the outcome of a development project by balancing focus areas by means of a set of indicators for a set of strategic objectives	[15-16]
BIKVA	Critical, subjective assessment of an existing practice	[17-18]
Clinical/diagnostic performance	Measurement of diagnostic 'correctness' (for instance measures of accuracy and precision) for IT-based expert systems and decision support systems	[53-55]
Cognitive assessment	Assessment of the cognitive aspects of the interaction between an IT-system and its users, for instance: ▫ Identification of where and why operation errors occur ▫ Identification of areas for a focused improvement of the user friendliness	[51-52, 56]
Cognitive Walkthrough	Assessment of user 'friendliness' on the basis of a systems design, from specifications, muck-ups or prototypes of the system, aiming at judging how well the system complies with the users' way of thinking, for instance: ▫ Identification of where and why operation errors occur ▫ Identification of causes behind problems with respect to the user friendliness and consequently identify areas for improvement	[57-58]
Delphi	1) (Qualitative) assessment of an effect, for instance where the solution space otherwise is too big for handling 2) Exploration of development trends 3) Elucidation of a problem area, for instance prior to a strategic planning	[19-22]
Effect assessment	Measurement of the effect, i.e. the consequence or impact in its broad sense) of an IT-based solution, - with or without the original objective as a frame of reference	[60-64]
Field study	Observation of and within an organisation to identify its practice and to elucidate mechanisms controlling change	[23-24]
Focus Group Interview	Is in principle used for the same purposes as other interview methods. In practice, the method is most relevant during the early analysis phase, for instance where attitudes or problems of social groups needs elucidation, or when a model solution is being established	[25-28]

Functionality assessment	1) Validation of objectives fulfilment (realisation of objectives), i.e. the degree of compliance between the desired effect and the materialised solution 2) Effect assessment (also called: impact assessment) 3) Identification of problems within the relation between work procedures and the IT system's functional solution The method will capture severe ergonomic and cognitive problems, but is not dedicated at capturing details of this type	[10]
Grounded Theory	Supportive analytical method for data acquisition methods that generate textual data, for instance some open questionnaire methods and interviews (individual and group interviews)	[30-32]
Heuristic evaluation	Is used, when no other realisable possibilities exist, for instance when: ▫ The organisation does not have the necessary time or expertise ▫ There are no formalised methods ▫ There is not yet something tangible to assess	[33]
Interviews (non-standardised)	Is in particular suited for elucidation of individuals' opinions, attitudes and perception regarding phenomena and observations	[34-36]
KUBI	Optimisation of the outcome of a long-term development project, based on a set of user or customer/client defined value norms and objectives	[18]
Prospective time studies	Measurement of development trends, including the effect of an initiative	[65-68], see also [1]
Questionnaires (non-standardised)	Questionnaires are used to answer a wide range of questions, but for investigations requiring a high level of accuracy, its main application area is (qualitative) investigations of subjective aspects	[39-43]
RCT, Randomised controlled studies	Verification of efficacy, i.e. that the IT-system – under ideal conditions – makes a difference for the patient care. In particular used for studies of decision-support systems and expert systems	[55, 59]
Risk assessment	Identification and subsequent monitoring of risk factors, enabling the possibility of taking actions before an unfortunate situation develops	[44]
Social network analysis	Assessment of relations among elements within an organisation (such as individuals, professions, departments or other organisations), which influence the acceptance and use of an IT-based solution	[45]
Stakeholder analysis	Assessment of stakeholders' features and their inner dynamics, aiming at the identification of participants for the accomplishment of a given task, problem solving or project	[37, 46]
SWOT	Situation analysis: establishment of a holistic view of a situation or a model solution	[1]
Technical verification	Verification that the agreed functions are present, work correct and in compliance with the agreement. This may take place for instance in connection with delivery of an IT-system or prior to daily operation, and at any subsequent change of the IT-system (releases, versions, and patches)	[1]

Usability	Assessment of user friendliness in terms of ergonomic and cognitive aspects of the interaction (dialogue) between an IT-system and its users	[47-50]
User acceptance and satisfaction	Assessment of users' opinion, attitudes, and perception of an IT-system at daily operation	[1]
Video recording	Monitoring & documentation as a means for analysis of what/how the work procedures respectively the users' activities are actually carried out, or for investigation of complex patterns of interaction	[51-52]
WHO: Framework for the assessment of strategies	Assessment of different (development) strategies either individually or as a comparative analysis	[1]

4.5 Generally useful information

There is a number of information that cannot be categorised under 'methods', but which anyway should be included, because the understanding of these issues is valuable. In general, the aspects within the table below are valid for all phases within the lifecycle.

Information	**Application area**	**References**
Documentation in an accreditation situation	Planning of assessment activities in connection with purchase of a 'standard' IT-system, when the user organisation is or considers becoming certified or accredited	[69-70]
Measures and metrics	Metrics and measures are used throughout evaluation, irrespective of whether it is constructive or summative. Planning of an assessment / evaluation study includes the conversion of an evaluation purpose to specific measures, and subsequent establishing metrics for their measurement	[1, 71-72]
Standards	A number of deFacto and deJure standards exits, which each defines a series of issues like the contents of a User Requirements Specification, verification of an IT-system, quality aspects of an IT-system, as well as roles and relations between a user organisation and a vendor in connection with assessment	[1]

5. REFERENCE LIST

1. Brender J. Handbook of Methods in Technology Assessment of IT-based Solutions within Healthcare. Aalborg:EPJ-Observatoriet, ISBN: 87-91424-04-6, June 2004. 238pp (in Danish)

2. Ammenwerth E, Brender J, Nykänen P, Prokosch H-U, Rigby M, Talmon J. Visions and strategies to improve evaluation of health information systems - Reflections and lessons based on the HIS-EVAL workshop in Innsbruck. Int J Med Informatics 2004;**73(6)**:479-491.

3. Haimes YY, Schneiter C. Covey's Seven Habits and the Systems Approach: A Comparative Approach. IEEE transactions on Systems, Man and Cybernetics 1996;**26(4)**:483-487.

4. Symons V, Walsham G. Evaluation of Information Systems: A Social Perspective. In: In: Bullinger HJ, Protonotarios EN, Bouwhuis D, Reim F, eds. Information Technology for Organisational Systems, Concepts for Increased Competitiveness. Proceedings of the 1st European Conference - EURINFO '88. Amsterdam:North-Holland. 1988:204-211.

5. Moehr JR. Evaluation: salvation or nemesis of medical informatics? Comp Biol Med 2002;**32**:113-125.

6. Arbnor I, Bjerke B. Methodology for Creating Business Knowledge. Thousand Oaks:Sage Publications, 1997.

7. Brender J. Methodology for Constructive Assessment of IT-based Systems in an Organisational Context. Int. J. of Med. Informatics 1999;**56**:67-86.

8. Berg M. Implementing information systems in health care organizations: myths and challenges. Int J Med Informatics 2001;**64**:143-156.

9. Quality Management and Quality Assurance Standards - Part 3: Guidelines for the Application of ISO 9001 to the Development, Supply and Maintenance of Software, 1st ed., corrected and reprinted 1993-05-01. Genève:ISO. 1991;**ISO 9000-3-1991 (E)**.

10. Brender J. Methodology for Assessment of Medical IT-based Systems - in an Organisational Context. Amsterdam:IOS Press, Studies in Health Technology and Informatics 1997;**42**.

11. Richart RH. Evaluation of a medical data system. I: Collen MF, ed. Proceedings of Conference on Medical Information Systems. San Fransisco, 28-30th January 1970.

12. Anderson JG, Aydin CE, Jay SJ, eds. Evaluating Health Care Information Systems, Methods and Applications. Thousand Oaks: Sage Publications, 1994.

13. Fernandez AM, Schrogie JJ, Wilson WW, Nash DB. A review and description of a "Best Practice" technology assessment process. Best Practices and Benchmarking in Healthcare 1997;**2(6)**:240-253.

14. Brender J. Methodological and methodical perils and pitfalls within assessment studies performed on IT-based solutions in healthcare. Technical Report of Virtual Centre for Health Informatics, May 2003; **03-1**. (ISSN 1397-9507).

15. Kaplan R, Norton D. Strategic learning & the balanced scorecard. Strategy & Leadership 1996;**24(5)**:18.

16. Gordon D, Geiger G. Strategic management of an electronic patient record project using the balanced scorecard. J Healthcare Inf Manag 1999;13:113-123.

17. Dahler-Larsen P, Krogstrup HK. Nye veje i evaluering. Århus:Systime. 2003. (in Danish)

18. Krogstrup HK. Evalueringsmodeller. Århus:Systime. 2003. (in Danish)

19. Adler M, Ziglio E. Gazing into the Oracle - The Delphi Method and its Application to Social Policy and Public Health, Jessical Kingsley Publishers, London, 1996.

20. Brender J, McNair P, Nøhr C. Research Needs and Priorities in Health Informatics. Int J Med Informatics 2000;**58-59(1)**:257-289.

21. Linstone HA, Turoff M. The Delphi Method, Techniques and Applications, Reading (Massachusets):Addison-Wesley Publishing Company, 1975

22. Dalkey NC. The Delphi Method: An Experimental Study of Group Opinion - Prepared for United States Air Force Project. Santa Monica (CA):Rand, 1969.

23. Jorgensen DL. Participant Observation, A Methodology for Human Studies. Newbury Park: Sage Publications. Applied Social Research Methods Series 1989;**15**.

24. Baklien B. Evalueringsforskning for og om forvaltningen. In: Foss O, Mønnesland J, eds. Evaluering av offentlig virksomhet, metoder og vurderinger. NIBRs PLUSS-SERIE 4-2000. NIBR, 2000, pp 53-77. (in Norwegian)

25. Stewart DW, Shamdasani PN. Focus Groups: Theory and Practice. Newbury Park: Sage Publications. Applied Social Research Methods Series 1990;**20**.

26. Harrison MI. Diagnosing Organizations: Methods, Models, and Processes, 2nd ed. Thousand Oaks: Sage Publications. Applied Social Research Methods Series 1994;**8**.

27. Robson C. Real World Research, A Resource for Social Scientists and Practitioner-Researchers, 2.ed. Oxford: Blackwell Publishers Inc, 2002, pp 269-291.

28. http://www.icbl.hw.ac.uk/ltdi/cookbook/focus_groups/index.html *(last seen 03.12.2003)*

29. Müller J. Lecture Note on Future Workshop, Design and Implementation. Technical Report of Dept. of Development and Planning, Aalborg University. August 2002.

30. Allen G. The Use of Grounded Theory as a Research Method: warts & all. In: Remenyi D, Brown A, eds. Proceedings of the European Conference on Research Methodology for Business and Management Studies, Reading 20-21. marts 2003. Reading:MCIL, 2003, pp. 9-19.

31. Balint S. Grounded Research Methodology – A Users' View. In: Remenyi D, Brown A, eds. Proceedings of the European Conference on Research Methodology for Business and Management Studies, Reading 20-21. marts 2003. Reading:MCIL, 2003, pp. 35-42.

32. Cronholm S. Grounded Theory in use – a review of experiences. In: Remenyi D, ed. Proceedings of the European Conference on Research Methodology for Business and Management Studies, Reading 29-30. april 2002. Reading:MCIL, 2002, pp. 93-100.

33. http://jthom.best.vwh.net/usability/ *(last visited 10.12.2003)*

34. Fowler Jr FJ, Mangione TW. Standardized Survey Interviewing: Minimizing Interviewer-Related Error. Newbury Park: Sage Publications. Applied Social Research Methods Series 1990;**18**.

35. Robson C. Real World Research, A Resource for Social Scientists and Practitioner-Researchers, 2.ed. Oxford: Blackwell Publishers Inc, 2002, pp 269-291

36. Rosenthal R. Experimenter Effects in Behavioral Research, enlarged edition. New York:Irvington Publishers, Inc., 1976.

37. Handbook for Objectives-Oriented Planning. 2nd edition. Oslo:Norwegian Agency for Development Cooperation. 1992.

38. Mettler PH, Baumgartner T. Large-scale participatory co-shaping of technological developments, first experiments with Pardizipp. Futures 1998;**30(6)**:535-554.

39. Ives B, Olson MH, Baroudi JJ. The measurement of User Information Satisfaction. Communications of the ACM 1983;**26(19)**:785-793.

40. Aydin CE. Survey Methods for Assessing Social Impacts of Computers in Health Care Organizations. In: Anderson JG, Aydin CE, Jay SJ, eds. Evaluating Health Care Information Systems, Methods and Applications. Thousand Oaks: Sage Publications, 1994, pp. 69-115.

41. Harrison MI. Diagnosing Organizations: Methods, Models, and Processes, 2nd ed. Thousand Oaks: Sage Publications. Applied Social Research Methods Series 1994;**8**.

42. Murphy CA, Maynard M, Morgan G. Pretest and post-test attitudes of nursing personnel toward a patient care information system. Comput Nurs 1994;12:239-244.

43. Sleutel M, Guinn M. As good as it gets? Going online with a clinical information system. Computers in Nursing 1999;**17(4)**:181-185.

44. Hall EM. Managing Risk: Methods for Software Systems Development. Reading:Addison-Wesley; 1998.

45. Anderson JG. Evaluation in health informatics: social network analysis. Comp Biol Med 2002;**32**:179-193.

46. Simmonds J, Lovegrove I. Negotiating a research method's 'Conceptual Terrain': Lessons from a stakeholder analysis perspective on performance appraisal in universities and colleges. In: Remenyi D, ed. Proceedings of the European Conference on Research Methodology for Business and Management Studies, Reading 29-30. april 2002. Reading:MCIL, 2002, pp. 363-373.

47. Bastien JMC, Scapin DL. Ergonomic Criteria for the Evaluation of Human-Computer Interfaces. Technical Report, INRIA. 1993;**156**.

48. Preece J. Part VI, Interaction Design: Evaluation. In: Human-Computer Interaction. Wokingham: Addison-Wesley Publishing Company, 1994.

49. Helander MG, Landauer TK, Prabhu PV, eds. Handbook of Human Computer Interaction, 2nd ed. Amsterdam:Elsevier Science B.V., 1997.

50. Mayhew DJ. The Usability Engineering Lifecycle, A Practitioner's Handbook for User Interface Design. San Francisco:Morgan Kaufmann Publishers, 1999.

51. Kushniruk AW, Patel VL. Cognitive computer-based video analysis: its application in assessing the usability of medical systems. In: Greenes RA, Peterson HE, Protti DJ, editors. Medinfo'95.

Proceedings of the Eigth World Congress on Medical Informatics; 1995 July 23-27; Vanvouver, Canada. Edmonton: Healthcare Computing & Communications Canada Inc; 1995. p. 1566-9.

52. Kushniruk AW, Patel VL, Cimino JJ. Usability testing in medical informatics: cognitive approaches to evaluation of information systems and user interfaces. Proceedings of AMIA Annual Fall Symposium; 1997. p. 218-222.

53. Brender J, Talmon J, Egmont-Petersen M, McNair P. Measuring quality of medical knowledge. In: Barahona P, Veloso M, Bryant J, eds. MIE 94 Proceeding of the Twelfth International Congress of the European Federation for Medical Informatics, Lisbon 22-26th May 1994. 1994, pages 69-74.

54. Egmont-Petersen M, Talmon J, Brender J, McNair P. On the Quality of Neural Net Classifiers. AIM 1994;**6 (5)**:359-381.

55. Friedman CP, Wyatt JC. Evaluation Methods in Medical Informatics. New-York:Springer-Verlag; 1996.

56. Beuscart-Zéphir MC, Brender J, Beuscart R, Ménager-Depriester I. Cognitive Evaluation: How to Assess the Usability of Information Technology in Healthcare. Computer Methods and Programs in Biomedicine 1997;**54(1-2)**:19-28.

57. Kushniruk AW, Kaufman DR, Patel VL, Lévesque Y, Lottin P. Assessment of a computerized patient record system: A cognitive approach to evaluating medical technology. MD Computing 1996;**13(5)**:406-415.

58. Beuscart-Zephir MC, Watbled L, Carpentier AM, Degroisse M, Alao O. A rapid usability assessment methodology to support the choice of Clinical Information Systems: a case study. In: Kohane I, ed. Proceedings of the AMIA 2002 Symposium on Bio*medical Informatics: One Discipline. 9th-13th November, San Antonio, Texas, 2002, pp. 46-50.

59. Kuperman GJ, Teich JM, Tanasjevic MJ, Ma'luf N, Rittenberg E, Jha A, Fiskio J, Winkelman J, Bates DW. Improving response to critical laboratory results with automation: Results of a randomised controlled trial. JAMIA 1999;**6(6)**:512-522.

60. Mitchell E, Sullivan F. A descriptive feast but an evaluative famine: systematic review of published articles on primary care computing during 1980-97. BMJ 2001;**322**:279-282.

61. Friedman CP, Abbas UL. Is medical informatics a mature science? A review of measurement practice in outcome studies of clinical systems. Int J Med Informatics 2003;**69**:261-272

62. van Gennip EM, Bakker AR: Assessment of effects and costs of information systems. Int J Biomed Comput 1995;39:67-72.

63. Sullivan F, Mitchell E. Has general practitioner computing made a difference to patient care? A systematic review of published reports. BMJ 1995;**311**:848-852.

64. Kimberly JR, Minvielle E. The Quality Imperative, Measurement and Management of Quality in Healthcare. London:Imperial College Press. 2000.

65. Evans RS, Pestotnik SL, Classen DC, Clemmer TP, Weaver LK, Orme JF, Lloyd JF, Burke JP. A computer-assisted management program for antibiotics and other antiinfective agents. New Engl J Med 1998;**338(4):**232-238.

66. Pennefather PM, Tin W, Clarke MP, Fritz S, Hey EN. Bias due to incomplete follow up in a cohort study. Br J Ophthalmol 1999;**83**:643-645.

67. Sparrow JM, Thompson JR. Bias: adding to the uncertainty, editorial. Br J Ophthalmol 1999;**83**:637-638.

68. Suissa S. The case-time-control-design: further assumptions and conditions. Epidemiology 1998;**9(4)**:441-445.

69. ISO/IEC. Informationsteknologi – Åbne Systemers sammenkobling – Overensstemmelses-prøvning, Del 5: Krav til prøvningslaboratorier og brugere i vurderingsprocessen. Hellerup:Dansk Standard. 1996;**ISO/IEC 9646-5**.

70. Danak. Retningslinie RL10, Anvendelse af edb i akkrediterede laboratorier, www.danak.dk, 2002. *(last visited 1.4.2003)*

71. Smith AE, Nugent CD, McClean SI. Evaluation of inherent performance of intelligent medical decision support systems: utilizing neural networks as an example. Int J Med Informatics 2003;**27**:1-27.

72. IEEE, Software Engineering Standards Committee. FP-03 Software Product Quality. http://standards.computer.org/sesc/sesc_pols/FP-03_Software_Product_Quality.htm. 2002 *(last visited 14.11.2002)*