PLANNING TO TEST SUCCESS – EMPLOYING THE GUIDELINES FOR GOOD EVALUATION PRACTICE IN HEALTH INFORMATICS (GEP-HI)

Michael Rigby
Professor of Health Information Strategy, School of Public Policy and Professional Practice, Keele University, Chancellor's Building, Keele University, Keele, Staffordshire, ST5 5BG, United Kingdom

Jan Talmon
School for Prevention and Primary Care – CAPHRI, Maastricht University, Maastricht, The Netherlands

Elske Ammenwerth
Institute for Health Information Systems
UMIT – University for Health Sciences, Medical Informatics and Technology, Hall in Tyrol, Austria

Jytte Brender
Department of Health Science and Technology
Aalborg University, Aalborg, Denmark

Nicolette de Keizer[6]
Department of Medical Informatics
Academic Medical Center, Amsterdam, The Netherlands

Pirkko Nykänen
Department of Computer Sciences
University of Tampere, Tampere, Finland

ABSTRACT

E-health is an important support technology in healthcare with considerable aspirations, but one which bypasses the normal requirement to prove a system's effectiveness or safety. There is therefore an ethical duty to evaluate systems scientifically, and thereby also to raise the scientific standing of the health informatics discipline. To assist in the development of this science, and arising out of a European Science Foundation exploratory workshop, guidelines on conducting robust evaluations have been prepared. This paper reports the results of that development work.

KEYWORDS

Health informatics; evaluation; guidelines; knowledge base

1. INTRODUCTION

Health informatics and e-health is a domain which has high objectives and ideals to improve the delivery of healthcare, and thus the health of the population. Indeed, it may be more aspiration than other domains such as the pharmaceutical sector, or development of health prostheses, where the goals tend to be more granular and relate to individual types of health need. However, and paradoxically, there is far less evidence to support these e-health hopes and claims. Unlike any other intervention in health care, informatics systems can be deployed without supportive proof either of benefits, or of absence of adverse effects.

In fact, it is known that there are anxieties about individual systems, while clinicians and others whose working practices have radical changes imposed on their working methods and practice are anxious over

their loss of established work patterns. Indeed, in some cases health informatics systems have been shown to have fatal outcomes (EFMI Working Group for Assessment of Health Information Systems, 2009). To remedy the avoidance of clinical or equivalent trials, and to build up an evidence base of proven benefits and other knowledge, evaluation of e-health needs to be rigorous, systematically critical and evidence-seeking. This paper introduces an important tool for that purpose, and argues for evaluation as an ethical imperative.

2. THE CONCEPT OF EVALUATION

To analyze this unmet need for rigorous evidence of benefits, a working conference on Health Informatics Systems Evaluation (HIS-EVAL) was organised by a small group of concerned individuals, and funded by the European Science Foundation (ESF). This resulted in the Declaration of Innsbruck, which summarizes the importance of evaluation in these words: "Health information systems are intended to improve the functioning of health professionals and organizations in managing health and delivering health care. Given the significance of this type of intervention, and the intended beneficial effect on patients and professionals, it is morally imperative to ensure that the optimum results are achieved and any unanticipated outcomes identified. The necessary process is evaluation and this should be considered an essential adjunct to design and implementation of health information systems" (Ammenwerth *et al,* 2004). Evaluation is the scientific means to assess the quality, value, effects and impacts of information technology (IT) in the health care environment. It is based on the concept of evaluation, defined as the "act of measuring or exploring properties of a health informatics application (in planning, development, implementation, or operation), the result of which informs a decision to be made concerning that system in a specific context" (Ammenwerth *et al,* 2004).

Given the known hazards and problems related to the use of IT in health care, and the paucity of impartial benefits and outcomes data, such evaluation is arguably not only a moral imperative, but also a logical need in order to ensure effective and full use by health professionals and indeed patients (Ammenwerth & Shaw 2005). However, it is known that there is a reluctance to undertake rigorous evaluation, with reasons ranging from vested policy or commercial interest to reluctance to divert 'treatment' or 'development' money to it, even though in the long term benefits from knowledge gained should significantly outweigh the initial investment (Rigby, 2001). And it is also known that e-health systems can have adverse outcomes, making monitoring and evaluation imperative (Gell, 2001; Ash *et al*, 2004; Berger, 2004.

3. THE DEVELOPMENT OF A STRUCTURED APPROACH

Reflective deliberations at the HIS-EVAL workshop identified the need to develop health informatics evaluation guidelines with authoritative standing. Following that workshop, two complementary guidelines have been developed, one on reporting standards and one on methodology.

The Statement on Reporting of Evaluation Studies in Health Informatics (STARE-HI) has been though a rigorous iteration and review process. It has been endorsed by the International Medical Informatics Association (IMIA) and by the European Federation of Medical Informatics (EFMI), and has been published (Talmon *et al.*, 2009). That statement indicates the elements which should be included in a robust report on a health informatics evaluation, if it is to be seen as scientific and credible. Such common standards are at the heart of modern evidence-based medicine, as exemplified for instance by the CONSORT standard for reporting randomised control trials (Begg *et al*, 1996). Given that health informatics and e-health systems affect patient care not only directly by intention, but also indirectly by affecting professional practice, there is no reason for e-health evaluations to be any less rigorous in their reporting, if the embedded evidence is to be accessible and effective.

The second element, which is the subject of this paper, are the Guidelines for Good Evaluation Practices in Health Informatics (GEP-HI) – the guidance on good practices for planning and execution of such evaluations, and in order to achieve the reporting standard.. Arguably, no informatics system should be implemented unless arrangements are put in place for an evaluation to prove benefits and learn lessons, even if this seems a diversion of creative resources (Rigby, 2001). The importance of applying evaluation as the partner to design and implementation, and the current poor state of methodology, has been emphasised at

international level (Brender 2006; Talmon, 2006; Rigby, 2006), hence the importance of developing these guidelines. This is work in hand, currently nearing completion, and can be viewed on line (Nykänen and Brender, 2008).

4. THE NEED FOR EVALUATION PRACTICE GUIDELINES

Given the general difficulty of conducting trials in health informatics, most studies are of a case study nature. However, as has been pointed out, "Case studies on evaluation are often not sufficiently grounded in theory, and established evaluation methods are frequently poorly applied. Evaluators are often insufficiently trained to select methods from various disciplines and to apply and combine these adequately. The proper design of evaluation studies, the selection of a framework to be applied and of methods to be used is difficult." (Ammenwerth *et al,* 2004). The development of the GEP-HI guidelines is intended to remedy this deficit.

There are difficulties and pitfalls in conducting evaluation studies, but no single global approach or methodology exists that is valid for all evaluation studies in any context. Therefore guidelines are needed that give us advice on how to design and how to carry out evaluation studies, and the issues to consider during the different study phases. Development of the GEP-HI guidelines has drawn on a wide range of literature, such as but not limited to (Schalock 2001; Kaplan and Shaw 2004; Ammenwerth and de Keizer 2005; Vimarlund & Olve, 2005; Davidson 2005; Fink 2005; Brender 2006; Friedman and Wyatt 2006; Westbrook *et al*., 2007; Hyppönen *et al*., 2007; Yusof *et al*., 2008a,b; Talmon *et al*., 2009).

5. THE ESSENCE OF THE GEP-HI GOOD EVALUATION GUIDELINES

The guidelines are divided into parts, corresponding to study phases ('phase' is used here in the sense defined by the International Standards Organisation (ISO) in ISO 9000-3 as "a segment of work"). The theoretical background for the phases is analogous to the general approach in information systems development models. They are further divided into tasks that from a planning perspective are coherent and meaningful components of the phases, and within each there are specific principles to be applied. Elements may appear in more that one phase, but the method and detail by which they are addressed may change according to the phase context.

5.1 The GEP-HI Study Phases

These phases are:

- 1. Study Exploration: the starting question of the evaluation study.
- 2. First Study Design: the preliminary design of the evaluation study.
- 3. Operationalisation of methods: making the design and evaluation methods concrete and compliant with the organizational setting and the information need, while taking into account the known pitfalls and perils. with sub-sections on **Error! Reference source not found.** and **Error! Reference source not found.**
- 4. Detailed Study Plan and Project Plan: providing plans, prescriptions and procedures detailed to the level necessary for the specific study.
- 5. Evaluation Study Implementation: activities related with the actual accomplishment of the designed evaluation study, with sub-sections on Project controlling and risk management - the good project management practices specifically for an evaluation study; and Reports and publications - how to report evaluation studies in terms of the STARE-HI guidelines.
- 6. Final Evaluation Phase: closing activities, such as archiving and accounting, and with a subsection on Report & Publications.

5.2 The GEP-HI Principles

The principles to be applied within each phase are shown in Table 1. The full details of these principles, with detailed rationale, are available on line (Nykänen P and Brender J, 2008).

Table 1. The GEP-HI Principles: Items recommended to be taken into account in any health informatics evaluation

PHASE	ITEM
1	**Study Exploration**
	1.1 The information need
	1.2 Primary audience
	1.3 Identification of the buyer / sponsor / study funding party
	1.4 The context of the evaluation study
	1.5 A first identification of stakeholders
	1.6 A first identification of (external) consultants
	1.7 A first sketch of the setting
	1.8 First exploration of evaluation methods to be used
	1.9 Exploring the restrictions of study execution and publication
	1.10 Budget
	1.11 Ethical, moral and legal issues
	1.12 Result of Study Exploration
	1.13 Formal accept to proceed to the next phase
2	**First Study Design**
	2.1 Elaboration of the rationale for the study
	2.2 Key evaluation issues/questions
	2.3 Budget
	2.4 Establishment of the design team
	2.5 Stakeholder analysis/Social Network analysis
	2.6 Study constraints
	2.7 Methods
	2.8 Organisational setting, the study context
	2.9 Technical setting, the study context
	2.10 Participants from the organisational setting
	2.11 Material and immaterial resources
	2.12 Time and timing
	2.13 Risk analysis
	2.14 Ethical, moral and legal issues
	2.15 Strategy for reporting
	2.16 Result of First Study Design
3	**Operationalisation of Methods**
	3.1 Study type
	3.2 Approach

PHASE	ITEM (continued)
	3.3 Assumptions
	3.4 Pitfalls and perils
	3.5 Skills
	3.6 Frame of reference
	3.7 Timing
	3.8 Justification
	3.9 Outcome measures
	3.10 Quality Control on data (measures)
	3.11 Participants
	3.12 Study flow
	3.13 Result of Operationalisation of Methods
	3.14 Ethical, moral and legal issues
4	**Detailed Study Plan and Project Plan**
	4.1 Project management
	4.2 Evaluation activity mapping
	4.3 Quality management
	4.4 Risk management
	4.5 Communication strategy
	4.6 Recruitment of necessary additional staff
	4.7 Result of Detailed Study Plan and Project Plan
5	**Evaluation Study Implementation**
	5.1 Establishment of the frame of reference
	5.2 Observation of changes
	5.3 Quality control of findings
	5.4 Interpretation of observations
	5.5 Continuous project management, quality management and risk management
	5.6 Regular reports
	5.7 Final result of Evaluation Study Implementation
6	**Final Evaluation Phase**
	6.1 Accountings
	6.2 Reports and publications
	6.3 Archiving
	6.4 Reporting guidelines
	6.5 Reporting scope
	6.6 The reporting message
	6.7 Authorship
	6.8. Ethical and moral aspects
	6.9 Preparation of reports / publications

6. DISCUSSION

Evaluation is a prerequisite justifying use of an e-health system, and should be shared for the benefit of the domain. Study reporting should be scientifically executed, compliant with the Statement of Reporting of Evaluation studies in Health Informatics (STARE-HI) (Talmon *et al.,* 2009). These Good Evaluation Practice for Health Informatics (GEP-HI) guidelines are intended to give assistance at the earliest possible point of an evaluation study. Based on a consensus-making process of evaluation experts they have been prepared to support future evaluation practices. The contribution of these guidelines to the health informatics domain is the build up of a robust scientific knowledge base on design and performance of evaluation studies. GEP-HI does not aim at providing detailed and prescriptive advice on what evaluation method to use in which setting for what purpose, as his is not feasible given the large variety in studies which is likely to be encountered.

The guidelines do, however, address issues that are relevant for successful planning and execution of a study in a form that is easy to follow yet without the details that would turn the guidelines into a rigid 'cook book'.

REFERENCES

Ammenwerth E, Brender J, Nykanen P, Prokosch HU, Rigby M, Talmon J., 2004. Visions and strategies to improve evaluation of health information systems. Reflections and lessons based on the HIS-EVAL workshop in Innsbruck. *Int J Med Inform*, 73 (6), 479-91

Ammenwerth E, de Keizer N, 2005. An inventory of evaluation studies of information technology in health care trends in evaluation research 1982-2002. *Methods Inf Med*; 44 (1):44-56.

Ammenwerth E and Shaw N, 2005. Bad health informatics can kill – Is evaluation the answer. *Meth Inf Med*; 44, 1-3.

Ash JS, Berg M, Coiera E. (2004). Some unintended consequences of information technology in health care: the nature of patient care information system-related errors. *J Am Med Inform Assoc.*;11(2):104-12.

Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, *et al*.(1996). Improving the quality of reporting of randomized controlled trials. The CONSORT statement, *Journal of the Americal Medical Association*;276(8):637-9.

Berger RG, Kichak JP (2004). Computerized Physician Order Entry: Helpful or Harmful? *J Am Med Inform Assoc.*; 11(2):100-3.

Brender J, 2006. *Handbook of Evaluation Methods for Health Informatics*. New York: Academic Press.

Davidson EJ, 2005. *Evaluation Methodology Basics, the Nuts and Bolts of Sound Evaluation*. Sage Publications: Thousand Oaks.

EFMI Working Group for Assessment of Health Information Systems, 2009. *Bad Health Informatics Can Kill* http://iig.umit.at/efmi/ (click sidebar Bad Health Informatics). (accessed 6 January 2009).

Fink A, 2005. *Evaluation Fundamentals, Insights into the Outcomes, Effectiveness and Quality of Health Programs*. Sage Publications: Thousand Oaks.

Friedman CF, Wyatt J, 2006. *Evaluation methods in medical informatics*. Second edition. Springer Verlag: New York.

Gell G (2001). Side effects and responsibility of Medical Informatics. *Int J Med Inform.;*64(2-3):69-81.

Hyppönen H, Salmivalli L, Nykänen P, Pajukoski M, Ruotsalainen P, 2007 Testing a theoretical framework for interdisciplinary IS evaluation: The case of Finnish Electronic Prescription. *International Journal of Healthcare Technology and Management,* 8, nos 1-2, 42-65.

Kaplan B, Shaw N, 2004. Future directions in evaluations research: People, organisational and social issues. *Meth Inf Med*; 43, 215-231.

Nykänen P and Brender J, 2008. Guidelines for Good Evaluation Practices in Health Informatics; available on http://iig.umit.at/efmi/gep-hi/GEP_HI_v014_July23_2008.doc (viewed 26 January 2009).

Rigby M, 2001. Evaluation: 16 Powerful Reasons Why Not to Do It - And 6 Over-Riding Imperatives. In: Patel V, Rogers R, Haux R, editors. *MEDINFO 2001: Proceedings*; IOS Press; 1198-1202.

Rigby M, 2006. Evaluation – the Cinderella Science of ICT in Health. In Kulikowski C, Haux R (Eds.) *IMIA Year Book of Medical Informatics 2006*; Schattauer, Stuttgart, 114-120.

Schalock RL. 2001. *Outcome-Based Evaluation*, 2nd Edition. Kluwer Academic / Plenum Publishers: New York, 2001.

Talmon J, 2006. Evaluation and implementation: A call for action. In Kulikowski C, Haux R (Eds.) *IMIA Year Book of Medical Informatics 2006*; Schattauer, Stuttgart, 11-15.

Talmon J, Ammenwerth E, Brender J, de Keizer N, Nykänen P, Rigby M, 2009. STARE-HI—Statement on Reporting of Evaluation Studies in Health Informatics; *Int J Med Inform*, 1-9.

Vimarlund V, Olve N-G, 2005. Economic analyses for ICT in elderly healthcare: Questions and challenges. *Health Informatics Journal*; 11 (4): 309-321.

Westbrook JI, Braithwaite J, Georgiou A, Ampt A, Creswick N, Coiera E, Iedema R, 2007. Multimethod evaluation of information and communication technologies in health in the context of wicked problems and sociotechnical theory. *J Am Med Inform Assoc*; (14): 746-755.

Yusof MM, Papazafeiropoulou A, Paul RJ, Stergioulas K, 2008a. Investigating evaluation frameworks for health information systems. *Int J Med Inform*; 77: 377-385.

Yusof MM, Kuljis J, Papazafeiropoulou A, Stergioulas K, 2008b. An evaluation framework for health information systems: Human, organisation and technology-fit factors. *Int J Med Inform*; 77: 386-398.