Evidence-Based Health Informatics
E. Ammenwerth and M. Rigby (Eds.)

doi:10.3233/978-1-61499-635-4-14

Evidence-based Health Informatics and the Scientific Development of the Field

Jeremy C. WYATT[a,1]
[a]*Leadership chair in eHealth Research, Leeds University, UK*

Abstract. We define and discuss the nature of Evidence-based Health Informatics (EBHI), the kind of evidence health informatics researchers must generate to make EBHI a reality, and how we should grade such evidence. We propose adding principle-based evaluation studies to the list of common evaluation study types, and outline how to carry out such studies to generate evidence that will prove useful for establishing EBHI. The main purpose of a principle-based evaluation study is to test the impact on system acceptability, usage or effectiveness of a generalizable system design principle, so we also explore when during the system design process such principles are needed, and which disciplines are most promising as sources of design principles. We conclude with some challenges for EBHI, a list of the benefits of adopting this approach, and a test to ensure that we are advancing in the direction of science, as opposed to pseudoscience.

Keywords. Health informatics, evaluation methods, evidence-based health, professional organisation and administration.

1. Introduction: What is evidence-based health informatics, and why does it matter?

Recently, a national body asked for advice on how to improve the quality of patient data captured by electronic health records. After a reminder that data quality has several dimensions [1], I suggested some strategies that might work and should be easy to implement: audits of data quality with weekly feedback to users, alternative screen layouts or data entry widgets, adding pop-up definitions of data items, or making certain data items obligatory. Fortunately, they did not ask for evidence about the relative impact of each strategy nor for which types of data or users each strategy is most appropriate. They would certainly have asked for such evidence if their question had been about which drugs work in a named disease, and there is copious high quality evidence about drug effectiveness. However, there is scarcely any good quality evidence about how to improve data quality [2], despite this being a common question and one which we in health informatics should be uniquely qualified and able to answer.

This is a major criticism of health informatics as a profession: we have not yet assembled a robust evidence base to answer basic questions about common clinical

[1] Corresponding author: Prof. Jeremy Wyatt, Leadership chair in eHealth, Research, Faculty of Medicine and Health, Worsley Building, University of Leeds, Leeds, LS2 9JT UK, j.c.wyatt@leeds.ac.uk. From 1-1-2016: Director, Wessex Institute of Health, University of Southampton, Alpha House Enterprise Road, Chilworth, SO16 7NS.

information management problems. Instead, we rely on experts, untested theories, poorly understood principles or low grade evidence. In short, the prevailing approach of health informatics is unscientific, so we cannot reliably predict the impact of the strategies we use to build or improve information systems.

To remedy this problem we need what can be called "Evidence-based Health Informatics (EBHI)". This can be defined by analogy with Evidence-based Medicine [3], which means doctors using the results of well-designed research in intact humans (evidence) to guide their patient management decisions, rather than relying on advice from experts or reasoning from first principles like pathophysiology. This requires the medical profession to take responsibility for developing and curating this knowledge base [4], a task which is now undertaken by the Cochrane Collaboration[2].

By analogy, EBHI means that the people designing, developing and implementing health information systems should be able to rely on an explicit evidence base derived from rigorous studies on what makes systems clinically acceptable, safe and effective – not on basic science or experts alone (see Figure 1).

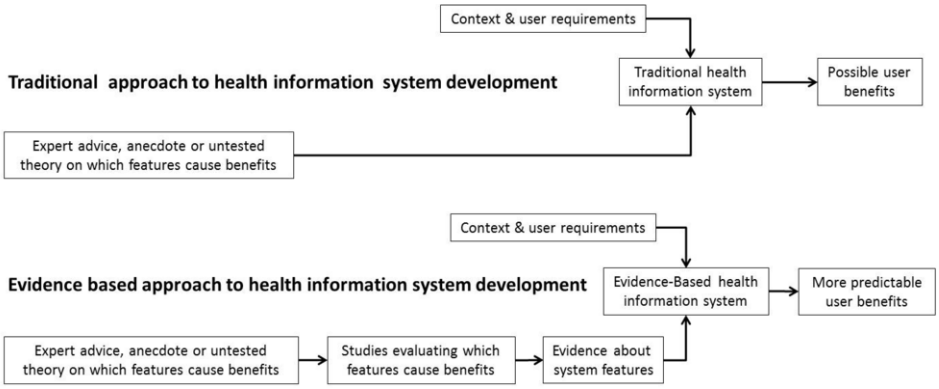

Figure 1. Comparison of traditional and evidence-based system development methods.

Once we in health informatics assemble this evidence base, this means that the design and implementation decisions taken by system developers will usually lead to predictably acceptable, safe, affordable and effective systems – which unfortunately is not the case at present [5]. The analogy for system development will be with cardiology or bridge building: with EBHI, system developers will become professionals relying on a proven body of knowledge (about test accuracy and drug effectiveness in the former case, or construction materials and how to use them in the latter), not craftsmen relying on a lifetime's experience of trial and error [6]. This will slow the excessive pace of technical innovation in our field, with every new technological development being tested for its contribution to important patient or health system outcomes. Over time, this evidence-based approach will lead to a number of benefits (see Table 1).

[2] For more information on the Cochrane Collaboration, see: Urquhart et al., Systematic reviews and meta-analysis of health IT. in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

Table 1. Some benefits of Evidence-Based Health Informatics for various stakeholder groups.

Stakeholder	Benefit
Patient	Safer, more effective health information systems; faster, more efficient care
Clinician / health professional	Systems that are easier to learn and use, fit better with clinical workflows, are safe and effective, with no surprises. Lower professional liability premiums as a result.
System developers	A clear set of guidelines for use in system development
Funders of clinical information systems e.g. health insurers	Systems that cost less and have predictable benefits
Tax payers, the public	Systems that cost less and have predictable and optimised benefits
Professional indemnity organisations	More reliable, effective health systems, so fewer legal claims against health professionals
System purchasers	A clear set of criteria to use during system procurement
People working in health informatics	Clarity about what to teach students Clarity about what works, when consulted about this A strong core of knowledge to inform future development of the profession
Regulatory organisations e.g. Medicine & Healthcare products Regulatory Agency (MHRA, UK), Food and Drug Administration (FDA, US)	An evidence base of tested principles against which to check new health information technologies
Clinical guideline developers	Good evidence on which to base their recommendations to use - or avoid - clinical information systems

The obvious next question is, what kind of evidence will we need to enable EBHI, and where will it come from? This is addressed next.

2. How to practise evidence-based health informatics?

2.1 What kind of evidence will we need to realise EBHI?

Evidence comes from primary and secondary research studies, but the best research design varies for each research question [7]. If we focus on the most important question in health informatics – which system design and development methods lead to safe and effective systems – then we can develop an approximate hierarchy of evidence for EBHI, analogous to the hierarchy of evidence for health technologies – drugs etc. At the top of this hierarchy are the most reliable sources of evidence, including systematic reviews[3] and randomized controlled trial[4] and the evidence gets steadily less reliable as we descend the hierarchy. A draft evidence hierarchy to support EBHI is shown in Table 2.

This implies an addition to the type of evaluation studies that we conduct in health informatics, adding to the usual studies (designed to answer the question "does it work") studies that ask "Will systems based on this generic design principle work better than other systems?". This is explored in the next two sections.

[3] See also: Urquhart et al., Systematic reviews and meta-analysis of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics 222, Stud Health Technol Inform, IOS Press, Amsterdam, 2016.

[4] See also: C.R. Weir, Ensuring the quality of evidence: Using the best design to answer health IT questions, in: ibid.

Table 2. Draft hierarchy of evidence to support EBHI, loosely based on [8].

Level	Type of evidence to support "What works?" questions
1a	Systematic reviews of well-designed impact studies designed to directly test a relevant design principle, with low heterogeneity
1b	Systematic reviews indirectly comparing well-designed impact studies that evaluate systems that demonstrate or lack a relevant design principle, with low heterogeneity
2	An individual randomised controlled study comparing the impact on real decisions or actions of a system designed according to a design principle or theory vs. a system not designed according to that principle
3a	Study comparing the safety or accuracy of a system based on the design principle against one not based on that principle, using real patient data
3b	Laboratory studies of simulated decisions or actions in response to a system based on the design principle vs. one not based on the principle, using real or simulated patient data
4	Untested theories or expert advice about what works in system design Anecdotes and case studies ("It worked for me")

2.2 How will this development change our evaluation methods?

Evaluation can be defined as carrying out studies to generate information to guide future decisions [9, chapter 1]. However, while all studies conform to this generic definition, from my observations over 35 years there are at least five different motives for conducting studies. These motives, along with some typical questions addressed by each type of study, are listed in table 3.

Table 3. Types of evaluation study.

Study type	Motive for carrying out study	Typical questions
1. Formative evaluation	How to improve an information system?	Is it accurate? Is it safe? Will people use it? How to improve it?
2. Summative evaluation	Can the finished system solve a specific problem?	Does this system work? How much does it cost? Will people use it?
3. Defensive evaluation	Was the funders' money spent well without making the situation worse?	Has anything improved since the system was implemented?
4. Self-interested evaluation	Can this study help the evaluator build their own CV?	Will this study have an impact on my colleagues?
5. Principle-based evaluation	Can this generic principle contribute to system design and EBHI?	Does this general design principle make systems more usable, effective, safer, less expensive, or more maintainable?

While the first four types of evaluation are relatively well known, the next section explains what we mean by the fifth.

2.3 How to design and conduct "principle-based evaluation"?

Principle-based evaluation means designing and conducting studies to test a generic design principle that if true, can guide future system development or implementation, thus helping to build the EBHI evidence base. Figure 2 below shows the steps that principle-based evaluation requires.

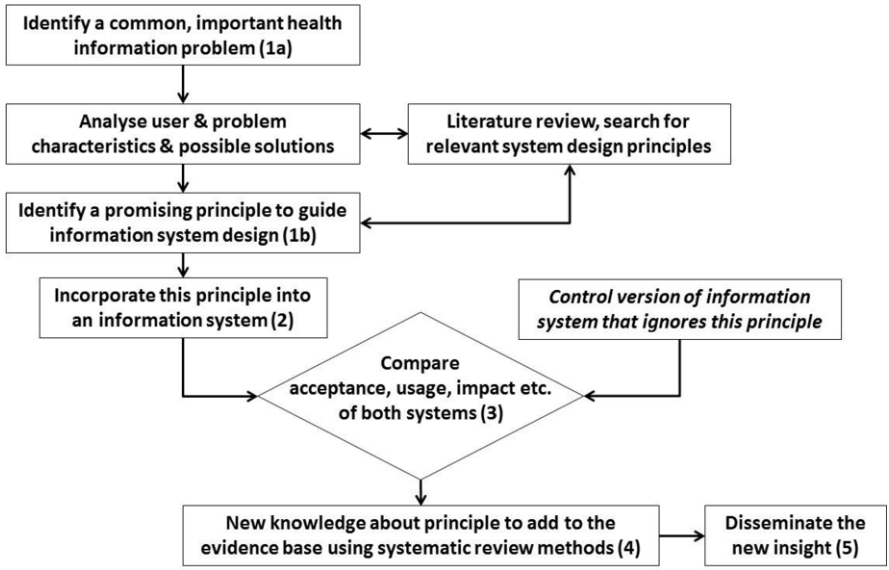

Figure 2. The process of designing and conducting principle-based evaluation studies (numbers refer to the text below).

If a researcher is planning a design principle-based evaluation study, they will need to carry out the following steps:

1. Working from a careful analysis of a common important problem (1a) in our domain (such as alert fatigue or poor data quality), identify a plausible generic system design principle or theory that may help resolve this (1b)
2. Use the selected principle to improve an existing information system, taking care that the only difference between the two systems is associated with application of the principle, not e.g. incorporating extra data or changing system usability (unless the principle concerns these specific actions)
3. Design one or more studies that rigorously test whether the design principle is supported or not, in terms of system acceptability, usability, accuracy, safety or impact on user decisions, actions or behaviours; or system maintainability or cost
4. If the study was small, integrate the results into the global evidence base of similar studies, using the well-established systematic review methods.
5. Accept and disseminate the results of their study, whatever these are – i.e. whether the principle makes sense to them or not. If the study was well designed, then its results should be respected.

There are some significant implications here for all evaluators. With the advent of EBHI, evaluators will need to think more clearly about their motives for carrying out a study and the consequences of this for their study design – particularly for the choice of controls. They will need to be clear about the differing aims of evaluation and their focus for each study. They will need to be familiar with a wide variety of evaluation methods, and how to identify and eliminate or control for biases and confounders [9, chapter 8]. They will also need to be aware of the obligation to publish their study

results (whether positive or negative), so that these are available to others aggregating evidence about what works and what does not.

2.4 Examples of studies and systematic reviews that contribute to EBHI

Some examples of studies that illustrate this approach and can potentially contribute to the health informatics evidence base are listed in table 4.

Table 4. Example studies and systematic reviews evaluating system design principles, in order of evidence grade.

Question	Type of study	Results	Source	Evidence grade (see Table 2) & comments
How to improve data quality?	Systematic review of 12 (mostly before-after) studies of various strategies in UK primary care	Most strategies appeared to have a positive effect, but study quality poor	Brouwer et.al. 2006 [2]	Evidence grade 1a. But systematic review was limited by poor study designs.
Does the use of psychological theory make a difference in behaviour change websites?	Systematic review and meta regression of 85 RCTs of theory based websites for health behaviour change	Use of theory to design website or recruit participants improved effectiveness by about one third of a standard deviation	Webb et.al. 2010 [10]	Evidence grade 1b. Use of theory may be confounded with better quality website design.
How much of a difference does tailoring and targeting make to text message impact?	Systematic review and meta regression of 19 RCTs of tailored SMS interventions for health behaviour change	Use of tailoring and targeting improves intervention effectiveness by 0.44 of a standard deviation	Head et.al. 2013 [11]	Evidence grade 1b. Use of tailoring may be confounded with better quality text design.
How to improve diagnostic accuracy?	RCT of a checklist	A well designed disease specific checklist improves accuracy by 10%	Adams et.al. 1986 [12]	Evidence grade 2. May reflect limited accuracy of junior doctors.
Can Fogg's principles of Persuasive computing improve websites for health-related decisions?	Online RCT of two websites to encourage 900 students to join NHS organ donation register	No – no difference (38% in both groups)	Nind et.al. 2009 [13]	Evidence grade 2. May only generalise to significant decisions such as organ donation.
Which kind of user interface speeds up data entry?	Experiment with 15 clinicians each entering 63 medical findings from 3 simulated cases using alternative prototype pen based user interfaces	Paged interface 5 seconds faster than scrolling. Complete list of codes 4 seconds faster than patient-specific list. Fixed position on screen 2 seconds faster than variable position.	Poon et.al. 1996 [14]	Evidence grade 3b. Limited to pen-based interfaces?
Can non-interruptive advice reduce	Within-subject experiment measuring	Prescribing alert in modal dialogue box twice as effective as	Scott et.al. 2011 [15]	Evidence grade 3b. Only tested one alert at a

errors?	prescribing errors in 20 junior doctors using case scenarios	same alert on ePrescribing system interface, but less acceptable.		time.

3. What kinds of design principle or theory to test?

So, what makes a good design principle to test? One aspect of this question is, *from which discipline or area* are promising design principles or theories likely to originate? To answer this question, consider a worked example: the design decisions faced by a team developing a typical eHealth system: an online forum to promote smoking cessation. Table 5 lists some of the fundamental decisions they need to take, together with possible disciplines or academic areas which could provide relevant design principles.

Table 5. Some design decisions made during the development of a sample information system, and possible origins of relevant design principles.

Design question / task	Discipline or area from which relevant design principles can originate
How to brand the website, and how to publicise it?	Marketing, public relations
What content to place on the website to attract smokers willing to quit?	Material to promote any of the techniques in the Behaviour Change Taxonomy [16]
How to encourage site visitors to enter, locate and retrieve information relevant to stopping smoking?	Search techniques; what makes risk / health information relevant Communication theory – common ground, etc. [17]
How to present information on the website in a manner that influences user decisions to quit?	Information design [18] Risk communication [19] Human decision making: heuristics and biases [20]
How to maximise the chances that a one-off user decision to stop smoking becomes a long term behaviour change?	Techniques drawn from the Behaviour Change taxonomy [16]

Another aspect of the big question is, what *kinds of design principle* are useful to test? Some properties of a candidate design principle that make a rigorous test valuable include that the design principle is:

- *Specific*: Sufficiently well formulated to be testable.
- *Actionable*: If proven, it would practically influence the design of health information systems.
- *Generic*: Can be applied across a range of information systems, user groups or contexts.
- *Credible*: The design principle appears well founded, so if proven is likely to be applied by others.
- *Enduring*: Such as theories about how people interact with and respond to information (eg. Risk perception), not theories about fleeting generations of technology (eg. High resolution vs. medium resolution virtual reality).
- *Novel or untested*: Not previously well tested for its impact on health information system design.

4. Some challenges arising from adopting the EBHI approach

Of course, the approach advocated will not solve every health informatics problem, such as use of the label "health informatics" by epidemiologists who can then attract funding intended for our discipline. As described earlier, we still need to conduct a wide range of evaluation studies more rigorously [9]. Also, we still need exploratory studies and experts to help us formulate plausible, generic design principles or theories for rigorous testing.

Another concern is that we should not over interpret the results of any individual study, as study results always vary randomly around the true effect size. So, we need to build evidence-based system design guidelines using systematic reviews[5], not on individual studies, unless we see "mega-trials" in our discipline as we see in cardiology, which in the current climate of health informatics evaluation scepticism seems highly unlikely. While it is tempting to use the systematic review method to compare the effectiveness of systems that do and do not incorporate a design principle from separate studies, caution is needed – which is why we consider such reviews as grade 1b evidence in Table 2. Using meta regression to test a design principle is not rigorous – all it shows is that there is an **association** between the principle and the outcome, not **causation**. To show causality, we need a direct randomised head-to-head comparison of the effectiveness of systems that did and did not incorporate the design principle in a single study (grade 2 evidence)[6], or ideally, a systematic review of head-to-head studies, which provides grade 1a evidence.

One dilemma is that while many design principles are generic (e.g. Schneiderman's user interface design guidelines [21]), some other principles (e.g. how to format displays of clinical data or alerts) may be bound up in the context of the specific users, data items or the task they support. The concept of ecological user interface design supports this: for each work domain or environment we design a user interface that supports this, with all the relevant information formatted in the optimum way to support the task in hand [22]. Realist approaches to evaluation and realist synthesis may have a place here to uncover what works, when, for whom and why [23][7].

5. Conclusions

In my opinion, the advent of principle-based EBHI marks the beginning of an exciting and fundamentally new approach to our discipline that, over time, will yield the evidence we need to place our discipline on a firmer base. It will allow us to authoritatively answer core questions fundamental to our discipline, such as "How to improve data quality?" with which this contribution started. This will bring greater confidence to our discipline and assure its ability to deliver safe, effective clinical

[5] See also: C. Urquhart et al., Systematic reviews and meta-analysis of health IT, in: E. Ammenwerth, M. Rigby (eds.), Evidence-Based Health Informatics, Stud Health Technol Inform 222, IOS Press, Amsterdam, 2016.

[6] See also: C.R. Weir, Ensuring the Quality of Evidence: Using the best design to answer health IT questions, in: ibid.

[7] See also: T. Otte Trojel et al., Going beyond systematic reviews: Realist and meta-narratives reviews, in: ibid.

information systems. However, it does mean more emphasis on rigorous study design and systematic reviews to identify and test potentially useful generic system design principles.

The benefits of a sound evidence base of system design principles or theories will include:

- The systems we produce will be reliably safe, efficient & predictable (like bridges).
- eHealth will evolve from an *intuitive craft* reliant on experts and apprenticeship into a *professional discipline*, making its decisions based on tested principles [6].
- There will be much less need for trial and error, or for re-invention of ad hoc systems that "seemed sensible at the time".
- Aspirational drives to 'modernise' or 'automate', followed by searches for available systems, will be considered inappropriate; instead there will be a call to grasp the proven benefits of validated systems.
- There will be no need to evaluate **every** version of **every** app, website, serious game etc., as long as the original one was built using tested principles, and the users or context of use have not changed too much to render these principles invalid.

A final comment is that to avoid what Grémy called "The idolatry of technology" (personal communication, Francois Grémy, 1999), health informatics should focus on science rather than on computer artefacts [18]. However, whenever we talk about science, we must also be wary of pseudoscience [24]. Fortunately, pseudoscience can be distinguished from science by the fact that scientific theories can be tested and ***disproved***, rather than confirmed [25]. So, health informatics professionals should avoid vague theories that cannot be tested, but also recognise that we will never know the limits of our new design principles until they fail us. However, meanwhile these design principles and theories will provide constructive new knowledge to inform future system design.

Recommended further readings

1. S.E. Straus, P. Glasziou, W.S. Richardson, R.B. Haynes, *Evidence-Based Medicine: How to Practice and Teach It*, 4th Edition, Churchill Livingston, Edinburgh, 2011.
2. C. Friedman, J. Wyatt, *Evaluation methods in biomedical informatics*, 2nd edition, Springer, New York, 2005.
3. C. Friedman, Where's the science in medical informatics? *J Am Med Inform Assoc* **2**(1) (1995), 65-7.
4. H.A. Heathfield, J. Wyatt, The road to professionalism in medical informatics: a proposal for debate, *Methods Inf Med* **34**(5) (1995), 426-33.
5. J. Wyatt, Medical informatics, artefacts or science? *Methods Inf Med* **35**(3) (1996), 197-200.

Food for thought

1. What are some disadvantages of the evidence-based approach to a scientific discipline?
2. Clinicians tend to consider clinical and cost effectiveness as the key evaluation criteria for a health technology. What alternative metrics might a computer scientist or a public health physician wish to consider, to help broaden the EBHI knowledge base?
3. How might a specific system design principle improve effectiveness while worsening system maintainability or widening health inequalities, for example? How do we manage those trade-offs?
4. Will health informatics as a discipline ever amass sufficient evidence-based design principles to allow us to develop and implement information systems with no need to carry out laboratory or field studies of safety and effectiveness?

References

[1] J.C. Wyatt, F. Sullivan, ABC of Health Informatics 1: What is health information? *BMJ* **331** (2005), 566-8.

[2] H.J. Brouwer, P.J.E. Bindels, H.C. Van Weert, Data quality improvement in general practice, *Fam Practice* **23** (2006), 529-36.

[3] D.L. Sackett, W.M. Rosenberg, J.A. Gray, R.B. Haynes, W.S. Richardson, Evidence based medicine: what it is and what it isn't, *BMJ* **312**(7023) (1996), 71-2.

[4] J. Wyatt. Use and sources of medical knowledge. *Lancet* **338** (1991), 1368-1373.

[5] B. Middleton, M. Bloomrosen, M.A. Dente, B. Hashmat, R. Koppel, J.M. Overhage, T.H. Payne, S.T. Rosenbloom, C. Weaver, J. Zhang, American Medical Informatics Association: Enhancing patent safety and quality of care by improving the usability of electronic health record systems: recommendations from AMIA, *J Am Med Inform Assoc* **20**(e1) (2013), e2-8.

[6] H.A. Heathfield, J. Wyatt, The road to professionalism in medical informatics: a proposal for debate, *Methods Inf Med* **34**(5) (1995), 426-33.

[7] D.L. Sackett, J.E. Wennberg, Choosing the best research design for each question, *BMJ* **315**(7123) (1997), 1636.

[8] D.L. Sackett, Screening in family practice: prevention, levels of evidence, and the pitfalls of common sense, *J Fam Pract* **24**(3) (1987), 233-4.

[9] C. Friedman, J. Wyatt, *Evaluation methods in biomedical informatics*, 2nd edition, Springer, New York, 2005.

[10] T.L. Webb, J. Joseph, L. Yardley, S. Michie, Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy, *J Med Internet Res* **12**(1) (2010), e4.

[11] K.J. Head, S.M. Noar, N.T. Iannarino, N.G. Harrington, Efficacy of text messaging-based interventions for health promotion: a meta-analysis, *Soc Sci Med* **97** (2013), 41-8.

[12] I.D. Adams, M. Chan, P.C. Clifford, W.M. Cooke, V. Dallos, F.T. de Dombal, M.H. Edwards, D.M. Hancock, D.J. Hewett, N. McIntyre, et al., Computer aided diagnosis of acute abdominal pain: a multicentre study, *Br Med J (Clin Res Ed)* **293**(6550) (1986), 800-4.

[13] T. Nind, J. Wyatt, I. Ricketts, P. McPate, J. Liu, *Website credibility and intervention effectiveness*, in Proceedings of the Persuasive Technology and Digital Behaviour Intervention Symposium: a symposium at the AISB 2009 Convention (6-9 April 2009) Heriot-Watt University, Edinburgh, Scotland, Society for the Study of Artificial Intelligence and Simulation of Behaviour, 2009, pp. 36-39.

[14] A.D. Poon, L.M. Fagan, E.H. Shortliffe, The PEN-Ivory project: exploring user-interface design for the selection of items from large controlled vocabularies of medicine, *J Am Med Inform Assoc* **3**(2) (1996), 168-83.

[15] G.P. Scott, P. Shah, J.C. Wyatt, B. Makubate, F.W. Cross, Making electronic prescribing alerts more effective: scenario-based experimental study in junior doctors, *J Am Med Inform Assoc* **18**(6) (2011), 789-98.

[16] S. Michie, N. Hyder, A. Walia, R. West, Development of a taxonomy of behaviour change techniques used in individual behavioural support for smoking cessation, *Addict Behav* **36**(4) (2011), 315-9.
[17] E. Coiera, When conversation is better than computation, *J Am Med Inform Assoc* **7**(3) (2000), 277-86.
[18] J.C. Wyatt, P. Wright, Medical Records 1: Design should help use of patient data, *Lancet* **352 (**1998), 1375-8.
[19] G. Gigerenzer, A. Edwards, Simple tools for understanding risks: from innumeracy to insight, *BMJ* **327**(7417) (2003), 741-4.
[20] A. Tversky , D. Kahneman, The framing of decisions and the psychology of choice, *Science* **211** (4481) (1981), 453-8.
[21] B. Schneiderman, C. Plaisant, *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 5th edition, Pearson press 2009.
[22] *Ecological interface design*, https://en.wikipedia.org/wiki/Ecological_interface_design, last access 11 February 2016.
[23] R. Pawson, J. Greenhalgh, C. Brennan, E. Glidewell, Do reviews of healthcare interventions teach us how to improve healthcare systems? *Soc Sci Med* **114 (**2014), 129-37.
[24] S.O. Hansson, "Science and Pseudo-Science", *The Stanford Encyclopedia of Philosophy* (Spring 2015 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/spr2015/entries/pseudo-science, last access 11 February 2016.
[25] J. Wyatt, Medical informatics, artefacts or science? *Methods Inf Med* **35**(3) (1996), 197-200.