ELSEVIER

International Journal of Medical Informatics 64 (2001) 15–37

International Journal of Medical Informatics

www.elsevier.com/locate/ijmedinf

Evaluating informatics applications—clinical decision support systems literature review

Bonnie Kaplan *

Center for Medical Informatics, Yale University School of Medicine, New Haven, CT, USA

Received 24 September 2000; accepted 5 July 2001

Abstract

This paper reviews clinical decision support systems (CDSS) literature, with a focus on evaluation. The literature indicates a general consensus that clinical decision support systems are thought to have the potential to improve care. Evidence is more equivocal for guidelines and for systems to aid physicians with diagnosis. There also is general consensus that a variety of systems are little used despite demonstrated or potential benefits. In the evaluation literature, the main emphasis is on how clinical performance changes. Most studies use an experimental or randomized controlled clinical trials design (RCT) to assess system performance or to focus on changes in clinical performance that could affect patient care. Few studies involve field tests of a CDSS and almost none use a naturalistic design in routine clinical settings with real patients. In addition, there is little theoretical discussion, although papers are permeated by a rationalist perspective that excludes contextual issues related to how and why systems are used. The studies mostly concern physicians rather than other clinicians. Further, CDSS evaluation studies appear to be insulated from evaluations of other informatics applications. Consequently, there is a lack of information useful for understanding why CDSSs may or may not be effective, resulting in making less informed decisions about these technologies and, by extension, other medical informatics applications. © 2001 Published by Elsevier Science Ireland Ltd.

Keywords: Evaluation; Decision support; CDSS; Clinical decision support systems; Clinical practice guidelines; Randomized controlled clinical trials

1. Introduction

Systems to aid in medical decision making were introduced over 25 years ago. Relatively few are in general use in clinical settings. Despite their potential usefulness, the lack of widespread clinical acceptance long has been of concern among researchers and medical informaticians [1–3].

This paper reviews literature that focuses on evaluation of clinical decision support systems (CDSS). The paper discusses the following key findings: The main emphasis is on changes in clinical performance that could

* Kaplan Associates, 59 Morris Street, Hamden, CT 06517, USA. Tel.: +1-203-777-9089; fax: +1-203-777-9089.
E-mail address: bonnie.kaplan@yale.edu (B. Kaplan).

1386-5056/01/$ - see front matter © 2001 Published by Elsevier Science Ireland Ltd.
PII: S1386-5056(01)00183-6

affect patient care. Many evaluations of CDSSs use designs based on laboratory experiment or randomized controlled clinical trials (RCTs) to establish how well the systems or physicians perform under controlled conditions. Other approaches to evaluation, such as ethnographic field studies, simulation, usability testing, cognitive studies, record and playback techniques, and sociotechnical analyses rarely appear in this literature. As was the case over ten years ago, few systems have been evaluated using naturalistic designs to study actual routine CDSS use in clinical settings. Consequently, the CDSS evaluation literature focuses on performance or specific changes in clinical practice patterns under pre-defined conditions, but seems lacking in studies employing methodologies that could indicate reasons for why clinicians may or may not use CDSSs or change their practice behaviors. Further, there is little reference in the CDSS literature to a theoretical basis for understanding the many issues that arise in developing and implementing CDSSs. In addition, the studies concern physicians to the near exclusion of other clinicians or potential users. Lastly, the literature seems not to be informed by studies of other medical computer applications, such as hospital information systems (HISs), computer based patient records (CPRs), physician order entry (POE), or ancillary care systems. These studies could provide useful insights into issues that likely would be relevant to acceptance and use of CDSSs.

2. Literature review methods

An automated literature search was done using Medline with the assistance of a librarian. This search identified papers classified as about a: (1) decision support system; (2) clinical decision support system; (3) expert system; and (4) decision aid. 'CDSS' has a variety of definitions. Any system that was considered a CDSS by the authors and catalogers of the papers reviewed was considered so for purposes of this review. This decision was made, instead of using an a priori definition of CDSS, so as to provide a view of the literature as it is presented and categorized by those involved. Using the authors' and catalogers' keywords is indicative of how those authors wish to have their work categorized and how this work is viewed within the discipline. It indicates how 'CDSS' is construed by those who are working within or commenting upon this area. Moreover, an a priori definition could result both in excluding papers authors consider as reporting on CDSSs, and in biasing results towards some particular type of system or definition. Further, the focus here is on evaluation, not on any particular type of CDSS. Hence, as in the guide published by Journal of American Medical Association to using articles evaluating the clinical impact of a CDSS [4], the search did not focus on any particular types of CSSS, such as alerting systems or diagnostic systems, but included them all.

The automated search spanned the years 1997 and 1998. To supplement the automated search, a manual search also was done. This included papers that had been referenced frequently by other papers, papers and authors known by reputation, review papers, papers in recent journals and proceedings, and books. The manual search was not limited in time period, but included years both before and after the automated search. This was especially the case for well-known review papers. Including recent review papers provided a more comprehensive scope to this undertaking. By examining review papers and commentaries that were published in past years, current work could be compared with prior trends in the CDSS literature. Doing so also

helped insure that significant works over the history of CDSSs were considered. Inclusion criteria for the manual review were any work concerning evaluation or success factors for expert systems or clinical decision support systems, and works describing evaluations of other systems or of evaluation approaches.

Papers identified in the search, but that clearly were irrelevant, were omitted from further consideration, leaving over 140 items that were reviewed thoroughly. Of these, only ten were found to be totally irrelevant. Those that were reviewed included research reports, editorials, reviews, descriptive and normative writings—in short anything that Medline returns from a search—and books. What follows is an analysis of themes and trends in the literature that was reviewed.

3. Usefulness of CDSSS

The literature indicates a general consensus that clinical decision support systems are thought to have the potential to improve care, or at least to change physicians' behavior [5]. Reminders [6–10]. alerts [11–17], treatment plans [6], and patient education [6], have been reported as effective in changing practice behaviors. Evidence of positive effect is more equivocal for guidelines [18–21]. Some studies suggest that guidelines are effective [19,22–28], and others that they are not [19,29]. There have been substantial rates of physician noncompliance with standards [29,30]. There is little evidence that physicians comply with guidelines, whether or not incorporated into a CDSS [20,27,31–34]. Whether systems aid physicians with diagnosis also is unclear [8,35–38].

Some see these results as exciting validations of the value of CDSSs. Others point out that, at best, the results are a 'disappointment' [36]. In addition, although physicians' behavior may be shown to change, there has been little study of whether the thinking behind the behavior has changed [39]. Studies of patient outcomes showed little significant improvement [5]. It also has been difficult to establish that patient outcomes have been affected [8,20,29,40,41]. Lastly, there is general consensus that a variety of systems are little used despite their demonstrated or potential benefits [18,42–47].

4. Evaluations of CDSS

Appendix A profiles all the evaluation studies of CDSSs found in the literature search. There are 27 studies reported in 35 papers. Papers reporting related studies are counted as one study each, though they are listed separately. Two of the 35 papers [48] are substantially the same, and, therefore, listed as one entry in the table.

A review of the studies in Appendix A suggests several notable tendencies:

1. As Appendix A shows, most studies are of specific changes in clinical performance that could affect patient care.
2. As is evident from Appendix A, most studies use an experimental or RCT design. With only six multi-methods studies, plus three more using qualitative methods, methodological diversity is limited. Other approaches to evaluation [49,50], such as ethnography, simulation, usability testing, cognitive studies, record and playback techniques, network analysis, or sociotechnical analyses rarely appear in this literature. Few studies involve field tests of a CDSS and almost none (two studies of CDSSs per se [51,52]) use naturalistic

designs in actual clinical settings with real patients (although one study used simulated patient encounters with actors playing the part of patients [37], and a number of studies of effects of alerts or reminders are based on actual treatment records).

3. There is little theoretical discussion in these papers. One study presented a theoretical analytical model [53]. Although a few mention theory, explicit theory is absent from most papers. Tacitly, papers are permeated by a rationalist or rational choice perspective.
4. As indicated in Appendix A, studies concern physicians to the near exclusion of other clinicians or potential users such as patients, administrators, project team members, insurers etc. Three studies include nurses [48,51,54]; one included providers, assistants, and patients [55]; and one concerns the project team and project history [54,56,57].
5. Judging from citations as well as the text, there is little mention of evaluations of other informatics applications.

These trends are reflected in recent review papers as well, as shown in Appendix B, which summarizes these review papers.

Discussion of these tendencies follows, with focus on the first three. This paper describes and analyzes the literature. Fuller implications of these observations, together with an analytical critique of the literature, are discussed elsewhere in this volume [58].

4.1. Focus on system and clinical performance

It has been reported that evaluations of CDSSs tend to concern system accuracy rather than either how well clinicians perform when actually using these systems, or the impact of system use on clinical care [35,43,59,60].[1] Elsewhere, it was found that evaluations of diagnostic systems tend toward process measures concerning performance of the system's user [61]. Evaluations identified in this review tend towards two kinds. The first are studies assessing CDSS accuracy and performance. A recent review emphasizes system functionality [62,63], and studies of decision-support systems usually rate the objective validity of the knowledge base, for example, by measuring performance against some gold standard [60,64,65]. However, only one study [43,66] listed in Appendix A concerns system performance. [2]

Although few applications are evaluated in practice [67], the second kind of evaluation, which dominates in Appendix A, concerns patient care more directly. Appendix C lists studies according to the kind of CDSS involved. As shown in Appendix C, most of the evaluation studies (21 of 27 studies) concern systems for alerts or reminders (nine papers), guidelines (six studies), and diagnosis (six studies). These studies are of specific changes in clinical performance that could affect patient care. This preponderance also is evident in Appendix B. Some of the studies investigate

[1] It is possible this depends on when the observation was made (or, as suggested in the next footnote, on search criteria). There is some evidence to suggest, both in others' work as well as in this review, that there has been a shift from system performance to user behavior. However, these citations are from 1998 to 1999, suggesting that the question of shift in emphasis bears further investigation.

[2] This may be due to the search criteria. For example, searching on 'diagnosis, computer assisted' identified 14 papers from the years 1997–2000. Of these, 11 assessed system performance. For reasons explained above, because no particular kind of CDSS was to be favored in this review, neither 'diagnosis, computer assisted' nor any other was included as a term in the automated search for this paper. Possibly this orientation predominates more in some publication outlets than in others. As noted in Appendix D, a large proportion of the papers were published in American Medical Informatics Association outlets and J. Am. Med. Assoc., while almost all papers were published by journals based in the US, even though authors may not be US based.

changes in physicians' practices, such as whether alerts affect prescribing behavior. Others are studies of changes in workflow and order processing, for example, time from order to delivery. This focus is well suited to study through experiment, and RCT is the dominant influence on study design. These kinds of measures are proxies for the very difficult issue of determining system effect on patient outcomes.

4.2. Study design and setting

RCTs and other experimental approaches have a long tradition as the standards for research design in clinical medicine [60,68,69]. It is not surprising that this also is the case in medical informatics. Van der Loo's study of evaluations of health care information systems from 1974 to early 1994 examined 108 studies against a set of quality standards. Study designs were ranked so that randomized trials were considered the 'highest', while qualitative designs are not discussed. Although 50% of the studies concerning what he classified as diagnostic systems or as treatment systems used an RCT design, only six of all the 108 studies met the stringent standard of economic analysis combined with an RCT. Disappointing quality scores for many of the studies he reviewed led him to call for a more a rigorous approach [61]. A substantial body of opinion in medical informatics supports this view. In pleading for controlled trials in medical informatics, for example, Tierney et al. state in the American Informatics Association editorial [70]:

> Only by performing rigorous clinical studies can we define whether a new information system will help, result in no change, or make the problem worse.

The CDSS literature clearly reflects this opinion. Normative papers call for randomized controlled trials [5]. Physicians are advised to apply the same criteria to assessing evaluations of CDSSs as of drugs or any other intervention. [4]. As indicated in Appendix 1, most papers reporting evaluations were experiments done under controlled conditions, even when in natural settings, so there was little methodological diversity. Of the papers listed in Appendix 1, two use these designs involving surveys [26,34], while only one uses surveys without experimental or RCT design [17]. Only four are multi-method (e.g. combinations of surveys, interviews, or observations) [71–74], plus two more studies are not for CDSSs per se but primarily involve computer-based clinical records [54,55,57]. Only the six multi-method studies plus three others [51,52,64] use qualitative methods, for a total of nine in all. Some of these authors explicitly stated how valuable they found using multiple methods, perhaps feeling a need to address the dominance of experimental approaches in this way. Five reported getting useful information through interviews and observations that would guide systems development [52,64,71–73].

As shown in Appendix 2, the RCT emphasis dominates for CDSS review papers. The same appears true for papers reviewing clinical guidelines' effectiveness, educational strategies, or barriers (though a comprehensive search was not done for these papers). Despite reviewers' claims that 'the simple randomised trial cannot be regarded as the gold standard in behavioural research' [25], their reviews are limited to randomized trials and other experimental and statistical methods considered rigorous.

Authors make a distinction between showing that a CDSS works under laboratory conditions and showing that it works under clinical conditions. Some recommend a multi-stage evaluation process, with evaluating functionality in real-life situations and evaluating system impact as the last stages [65]. Some 95% of

systems never reach the stage of field evaluation [37]. Apparently, there has been little change in this number over the years. It comes from a review published 10 years (1987) prior that noted that '[a]pproximately 90% of all computerized medical expert systems have not been evaluated in clinical environments' [75]. A few years later, according to a 1990 report [76]:

> [O]nly about 10% of the many medical knowledge-based systems that have been described over the years have been tested in laboratory conditions, while even fewer have been exposed to clinical trials.

Appendix A lists few studies involving field tests. Thus, it seems that very few CDSSs have been independently evaluated in clinical environments (or that no one has recounted them). This remains the case even though calls repeatedly have been made to field test CDSSs so as to demonstrate that they work in patient care settings, and even though some of those calls are from researchers who have not conducted their studies in this way, e.g. [36,66,75,76].

Some authors further recommend that these field settings be remote from and relatively independent of system developers because 'study design needs to rest upon making sure that the reasons for success or failure are clear' and 'be broad enough to detect both intended and unintended effects' [77]. Some call for assessing systems in actual use, under routine conditions, and for understanding why the results of such assessments are as they turn out to be. Nevertheless, they say little about how to achieve this understanding, and further, they either propose, or actually carry out, evaluations based on a clinical trials or experimental models, e.g. [35,36,76]. Clinical trials, even in practice settings, are considered the 'obvious' approach [59].

As substantiated in the appendices, the evaluation focus is on how CDSSs affect clinical processes or outcomes [5]. In what is perhaps the closest simulation to a real patient visit, Ridderikhoff and van Herk use cases constructed from real patient data and an actor playing the patient. They also include observational data in their report [37]. Berg [51] and Kaplan et al. [52] each are unusual in reporting detailed naturalistic observational field studies of CDSSs in actual use with real patients under routine clinical conditions.

A review of evaluations of all medical informatics applications reported in the 1997. AMIA Proceedings found patterns similar to those reported here. Almost all of those evaluations were of CDSSs and the primary evaluation design was modelled on controlled trials. Generally, individual systems were evaluated against expert human performance, or subjects were given simulated patient cases so that their performance with and without an automated system was compared [78].

4.3. Theoretical orientation

Although few authors discuss theory, this review indicates a strong theoretical preference underlying most studies. As indicated above, most employ an experimental or RCT design and use solely quantitative data collection and data analysis methods. Thus, studies reflect an objectivist epistemological stance and quantitative methodological approach [49,50]. They evidence a rationalist or rational choice perspective and focus on measurable variances by comparing effects of system use with other circumstances [78–81].

This perspective often is tacit. As one example, it was reported in 1990 that decision

aids in medicine usually are evaluated by measuring the structure of the aid, the function of the aid, or the impact of the aid on users and patients. While 'impact … on users and patients' might seem to imply a different approach, instead, here it refers to effects of the system on process measures, such as accuracy, timing, and confidence of decisions; or effects of the system on outcome measures, such as patient morbidity and mortality, or cost per procedures [76]. A similar orientation is evident in three categories of reasons that are given for evaluation: ethical, legal, and intellectual. Despite the apparent breadth of these three categories, the focus is on measurable economic and technical factors. For example, legal evaluation, in this instance, includes how effective and how safe a system is, and how it might change resource use. Thus, even where authors recognize that 'the unusual properties of medical expert systems' make it necessary to modify randomized double-blinded controlled trial for field trials, their suggestions remain within a rationalist framework [76].

This tacit perspective also is apparent among other evaluators. Advocates of a systems approach that includes taking full account of 'medical, economic, technical, organisational and behavioural dimensions' when doing an economic evaluation [82], thereby subordinate these concerns to economic analysis. Some discuss user acceptance without mentioning cultural and sociologic factors [4], while others state that these factors need to be considered. Nevertheless, these authors, like those who do not mention such contextual factors [46], discuss acceptability in terms of user- and machine-machine interfaces, response time, and similar technical issues. Some limit discussion of user acceptance to the interface while emphasizing that the purpose of evaluation is safety through accuracy and adequacy of the domain knowledge [83,84]. When considering the usability and acceptance of the interface, subjective measures such as user questionnaires and expert review are not valued highly [84], even though physicians consider having tools that add value to the practice setting more valuable than usability [71]. A broader emphasis on user satisfaction, if discussed at all, is on developing generic satisfaction instruments and appropriate controls [5]. As these examples illustrate, the underlying approach fits an objectivist, rationalist philosophical orientation and design employing quantitative methods to measure variance, even if not explicitly acknowledged.

5. Conclusions

Despite calls for alternatives, or recommendations to select designs congruent with system development stage and different evaluation questions [49,50,65,67], RCTs remain the standard for evaluation approaches for CDSSs [85,86], making evaluation traditions for CDSSs similar to those for other computer information systems, whether or not they may be intended for use in health care. Most commonly, systems, whether medical or not, have been evaluated according to selected outcomes pertaining to features such as technical or economic factors at the expense of social, cultural, political, or work life issues [79,80,87]. RCT and other experimental designs are excellent for studying system performance or specific changes in clinical practice behaviors, but not well suited for investigating what influences whether systems are used. Consequently, some other evaluation approaches have been developed, including simulation, usability testing, cog-

nitive studies, record and playback techniques, ethnography, sociotechnical analyses, and social interactionism among them.

Commentary concerning implementation issues and barriers to system use are little different today from what has been reported over the past 50 years [2]. This may be partly because system evaluations often ignore issues concerning user acceptance or changes in work, an omission also evident in the literature that was reviewed for this paper. By focusing on pre-specified outcome measures, evaluations do not examine processes of actual system use during daily activities [88]. As a result, we have excellent studies that indicate decreases in medication errors with physician order entry [11] or when notified by pharmacists or radiology technicians about drug alerts [16], changes in physician prescribing behavior for at least 2 years after a study [22], and greater compliance with guidelines [20]. Yet we have not sufficiently studied why these were the results. Nor have we investigated reasons behind other, less encouraging, findings. We have little understanding of why, for example, physicians agreed with 96% of one system's recommendations but only followed 65% of them [31], why, in another study, there was an overall increase in compliance with guidelines but the compliance rate still was low [27]; or, in another, why there was an increase in compliance, except for three items [28]; or why only the same four of six groups of preventive practice were improved with either reminders that were computer generated or those that were manual, but all six groups improved with computer plus manual reminders [10]. Despite these improvements, another study indicates that there were no significant differences in complying with guidelines between physicians who received computerized reminders and those who did not [19]. What accounts for these differences? Elsewhere, individuals found their post-implementation experiences fell short of their expectations [72]. Why did this happen, and how much does it matter? Study designs did not address questions that allow deeper understanding of these findings, understanding that could indicate why different results obtain in different studies. Consequently, we cannot learn what to do that might improve the practices that these studies measure.

Other research approaches are little reflected in the CDSS evaluation literature. These omissions are impoverishing our understanding of CDSS as they might actually be used [58]. RCT-type studies are excellent for demonstrating whether a particular intervention has a pre-specified effect. Such studies of CDSSs are valuable. Nevertheless, they tell us little about whether clinicians will incorporate a particular CDSS into their practice routine and what might occur if they attempt to do so. Such studies cannot inform us as to why some systems are (or will be) used and others are not (or will not be), or why the same system may be useful in one setting but not in another. They do not indicate why a CDSS may or may be not effective. Different study designs answer different questions. A plurality of methodological approaches and research questions in evaluation is needed so as to broaden our understanding of clinical acceptance and use of informatics applications [58].

Acknowledgements

I am grateful to Dr Richard Spivack of the US National Institute of Standards and Technology for his invaluable assistance in the automated literature search.

Appendix A

CDSS Evaluation

Authors	System	Study design	Findings
Bates et al. 1998 [11]*	Drug Alerts POE	Comparison of medication errors before and after implementation, and also with and without team intervention	POE decreased rate of medication errors. Team intervention conferred no additional benefit over POE.
Bates et al., 1999 [12]*	Drug Alerts POE	Comparison of medication errors at different time periods	POE decreased rate of medication errors
Berner et al. [43]+	Dx DSS	Comparison of physicians' performance on constructed cases	Physicians' performed better on the easier cases and on the cases for which QMR could provide higher-quality information.
Berner et al., 1994 [66]+	Dx DSS	Comparison of programs' performance	No single computer program scored better than the others. The proportion of correct diagnoses ranged from 0.52 to 0.71, and the mean proportion of relevant diagnoses ranged from 0.19 to 0.37.
Berg, 1997 [51]	Dx DSS	Case studies in clinical settings	Actor-network theory is used to describe how system implementation changed both the system and work practices.
Bouaud et al., 1998 [31]	Guidelines	Measured physicians' agreement and compliance with guidelines	Clinicians agreed with 96% of the recommendations and followed one of the recommendations in 65% of cases.
Buchan et al., 1996 [22]	Guidelines	Comparisons of prescribing behavior	Participation was followed by a favorable change in clinical behavior which persisted for at least two years.
Friedman et al., 1999 [35]†	Dx DSS	Comparison of physicians' Dx using different systems in laboratory setting	DSS consultation modestly enhanced subjects' diagnostic reasoning.

Gadd et al., 1998 [71]	DSS interface	Comparison of perceptions of different prototype versions of the system through video observation, surveys, and interviews	Features that improve perceived usability were identified.
Gamm et al., 1998 [72]	Computer based patient record	Comparison of pre- and post-installation survey data. Also did interviews and observations.	Pre-installation, most respondents were moderately positive about the helpfulness and utility of computerization in their practice. Post-installation experience fell short of those expectations.
Jha et al., 1998 [13]*	Drug Alerts	Compare computer-based adverse drug event (ADE) monitor against chart review and voluntary report by nurses and pharmacists	The computer-based monitor identified fewer ADEs than did chart review but many more ADEs than did stimulated voluntary report.
Kaplan et al., 1997 [52]	Guidelines	Case study using observation and interviews concerning diagnostic and treatment guidelines in psychiatry	Design suggestions and user acceptance issues were identified.
Karlsson et al., 1997 [64]	DSS	Study how clinicians viewed using this way of accessing information through interviews using 'stimulated recall'.	The major uses of the system were for patient-specific support and continuing medical education. Three parameters-relevance, validity, and work were important.
Kuperman et al., 1999 [14]	Lab Alerts	Compare time to treatment with and without automatically paging the physician.	The automatic alerting system reduced the time until treatment was ordered.

Lauer et al., 2000 [53]	Patient scheduling	Case study assessing system against a priori model.	The model helps provide a theory-based understanding for collecting and reviewing users' reactions to, and acceptance or rejection of, a new technology or system.
Litzelman et al., 1993 [26]‡	Reminder	Prospective, randomized, controlled trial. Compared compliance with computer-generated reminders between 2 groups of physicians.	Compliance with computer-generated reminders was higher in the group that received printed reminders and also was required to indicate response to reminders than in the group not required to indicate response.
Litzelman and Tierney, 1996 [34]‡	Reminders	Survey of physicians.	55% of computer-generated reminders were not complied with. Of those, 23% were not applicable and 23% would be done at the next visit. Of those to be done at the next visit, the stated reason was 84% because of lack of time this visit.
Lobach and Hammond 1997 [27]	Guidelines	Controlled trial comparing physicians randomized to receive encounter form with or without computer generated guidelines.	Providing the guidelines resulted in a two-fold increase in clinician compliance with the guidelines, but the overall compliance rate still was low.
McDonald, 1975 [30]	Reminder	Controlled crossover design with actual patient visits.	Prospective reminders reduced errors.
Monane et al., 1998 [15]	Drug alerts	Compare changes in prescription against expected baseline rate of change.	A computerized drug utilization review database linked to a telepharmacy intervention improved prescribing patterns.
Moore, 1994 [73]	Stroke Dx and Management	Pre-use and post-use interviews and surveys and participant observation.	Participants were positive toward computers in general. Findings also include reactions of different medical personnel to the clinical use of this system and changes needed to make it more effective in the clinical context.

Morgan et al., 1998 [28]	Guidelines	Compare compliance with guidelines at baseline and at 1 year and 5 years after.	After several years, compliance with guidelines was still up, except for 3 items, which went below baseline
Overhage et al., 1996 [19]	Reminders	Compare physicians randomized to receive computer reminders with those who do not receive them.	No significant differences were found between the groups in their compliance with guidelines.
Raschke et al., 1998 [16]	Drug Alerts	Measure number of times physicians changed orders consistent with drug alerts.	Physicians changed orders as a result of being notified by the pharmacist or radiology technician who screened the alerts.
Ridderikhoff and van Herk, 1997 [37]	Dx DSS	Physicians used DSS for simulated patients (actors) who presented cases.	The doctors' diagnostic accuracy was 43%. The role of the doctor in computer-aided diagnostics remains open to debate.
Ruland, 1998, 1999 [48]	Patient preferences	Comparison of nurses with and without patient preferences obtained through paper-based elicitation tool.	When nurses had patient priority information, nurses' care priorities changed to be more consistent with a patient's. This also improved patients' preference achievement and physical functioning.
Safran et al., 1998 [55]	Computer based patient record that includes alerts and reminders, and e-mail	Cognitive evaluation based on comparing (1) model of interactions among team members, developed from observing and analyzing work patterns, together with interviews, and (2) characterizations of traditional primary care units.	The least structured communication methods are also the most heavily used: face-to-face, telephone, and electronic mail. Most of the providers who were observed appreciated alerts and reminders, although the more expert practitioners found them annoying because they considered them unnecessary.

Schriger et al., 1997 [20]	Guidelines	Comparison of compliance measures between baseline, after intervention, and after computer system removed.	Use of a computer-based system improved documentation, compliance with guidelines, and percentage of charges spent on indicated activities, while decreasing overall charges. The parameters returned to baseline when the computer system was removed.
Sicotte et al. [54]§	CPR	Multiple case study involving interviews, focus groups, observations, and secondary documented sources at 4 hospitals.	Profound misconceptions in achieving a tighter fit (synchronization) between care processes and information processes were the main problems. Nurses were reluctant to use the system because it imposed a new reality involving uniformity and predictability in though and behavior.
Sicotte et al., 1996 [56]§	Feedback	Compared clinical record information at baseline and after intervention.	Reports to physicians of individual and group profiles failed to modify the physicians' practice profiles.
Sicotte et al., 1998 [57]§	Computer based patient record	Multiple case study involving interviews, focus groups, observations, and secondary documented sources at hospitals.	The implementation failed, primarily due to obstructions to achieving tighter synchronization between the care and information processes.
Warner et al., 1998 [17]	Alerts	Questionnaire for physicians to rate usefulness of alerts received by pager or e-mail (according to their preference).	Internal Medicine clinicians thought the alerts were helpful clinically. They wanted to keep the system after the end of the study, and they would recommend it to colleagues.

Weaver, 1994 [74]	Dx DSS	Qualitative case study using interviews and documents	The study identified barriers to diffusion and organizational impact on doctors, patients, paraprofessionals and other effects of Weed's PKC, a guidance system intended for routine use with any patient. Social factors hindered the PKC's development and potentially beneficial effects.
Wolf et al., 1997 [38]§	Dx DSS	Comparison of physicians' case work-up first without system, then with it.	No significant main effects of the DSS were found for: (1) level of experience; (2) whether or not subjects indicated they would seek a diagnostic consultation before using the DSS; or (3) whether or not they found the DSS consultation in fact to be helpful in arriving at a diagnosis

*, +, †, §, ‡, Studies that appear to be related are marked with the same symbol and counted as one study.

Appendix B. Reviews of evaluations

CDSS evaluations

Authors	System	Number	Study design	Findings
Balas et al., 1996 [6]	Clinical information systems	100 trials, 98 articles	RCT	Provider prompt/reminder, computer-assisted treatment planner, interactive patient education/therapy, and patient prompt/reminder were significantly successful interventions.

Hunt et al., 1998 [8]	CDSS	68 trials	Controlled	Some benefit was found in 66% of the studies reviewed. The CDSSs can enhance clinical performance for drug dosing, preventive care, and other aspects of medical care, but not convincingly for diagnosis. The effects of CDSSs on patient outcomes have been insufficiently studied.
Johnston et al., 1994 [41]	CDSS	28 trials	Controlled trials	Some CDSSs can improve physician performance. Studies are needed to assess effects on cost and patient outcomes.
Miller, 1994 [46]	Dx, DSS	All		Diagnostic decision support systems have become an established component of medical technology.
Nøhr, 1994 [60]	Dx DSS	10		Studies were reviewed to determine whether they concerned structure, process, or outcomes. All focused on structure measures concerning (six compared system performance against a gold standard) though two evaluated issues related to the process of care.

Shea et al., 1996 [10]	Reminders	16 trials	RCT	The study compared computer-generated to manual reminders to both and to none. Computer reminders alone and manual reminders alone improved practices for the same four of the six groups of preventative practices, with the computer plus manual reminders improving practices for all six groups.
Shiffman et al., 1999 [63]	Guidelines	25 studies	RCT and time-series correlational studies	Authors analyze functionality of guideline systems. Guideline adherence and documentation improved in those studies that measured it.
[5] Guidelines evaluations				
Cabana et al., 1999 [32]	Barriers to guideline implementation	76 articles that included 120 surveys and 5 qualitative studies		Authors classified barriers into 7 general categories, which they discuss.
Davis, Taylor-Vaisey, 1997 [33]	Guideline implementation	N/A	RCTs and trials that objectively measured physicians' performance or health care outcomes	The findings of the literature review may be grouped into 2 broad areas: those exploring the variables affecting physicians' adoption of Clinical practice guidelines in a naturalistic manner, and those describing outcomes of trials of educational interventions to change physicians' behavior or health care outcomes.

Davis et al., 1995 [7]	Guideline implementation strategies	99 trials	RCT	Effective change strategies included reminders, patient-mediated interventions, outreach visits, opinion leaders, and multifaceted activities. Audit with feedback and educational materials were less effective, and formal continuing medical education conferences or activities, without enabling or practice-reinforcing strategies, had relatively little impact.
Grimshaw and Russell, 1993 [25]	Guidelines	59 evaluations	RCT and 'other robust designs' that allowed for statistical analysis	Explicit guidelines improve clinical practice although the size of the improvements in performance varies considerably.

Appendix C. CDSSs evaluated

Type of CDSS	Authors
Alerts and reminders	Bates et al., 1998 [11]*
	Bates et al., 1999 [12]*
Nine studies	Jha et al., 1998[13]*
	Kuperman et al., 1999[14]
	Litzelman et al., 1993 [26]‡
	Litzelman, Tierney, 1996 [34]‡

	McDonald, 1975 [30] Monane et al., 1998 [15] Overhage et al,. 1996 [19] Raschke et al., 1998 [16] Safran et al., 1998 [55] Warner et al., 1998 [17]
Diagnosis Six studies	Berner et al., 1999 [43]+ Berner et al., 1994 [66]+ Berg, 1997[51] Friedman et al., 1999 [35]† Moore, 1994 [73] Ridderikhoff and van Herk, 1997 [37] Weaver, 1994 [74] Wolf et al., 1997 [38]†
Guidelines Six studies	Bouaud et al., 1998 [31] Buchan et al., 1996 [22] Kaplan et al. ,1997 [52] Lobach and Hammond, 1997 [27] Morgan et al., 1998 [28] Schriger et al., 1997 [20]

*, +, †, ‡, Studies that appear to be related are marked with the same symbol and counted as one study.

Appendix D. Where evaluations were published

Where published CDSS studies	#Papers	Design
AMIA proceedings	8	3 comparisons [28, 38†, 48] 4 qualitative methods [52,64,71,72] 1 questionnaire [17,31]
JAMIA	6	5 comparisons [12]* [13]*[14] [43]$^{+}$ [48] 1 model-based case study [53]

J. Am. Med. Assoc.	5	5 comparisons [11]*[15, 16, 20, 35†]
J Gen Int Med	2	2 comparisons including surveys [26]‡[34]‡
JMS	2	2 multi-method [54§, 57§]
NEJM	2	1 controlled cross-over [30]
Comparison [66]+		
AIM	1	comparison [55]
Am J Med	1	comparison [27]
Arch Int Med	1	comparison [19]
Fam Pract	1	comparison [22]
IJMI	1	1 simulation [37]
J. Beh. Med.	1	2 comparison [56]§
Books	3	3 qualitative methods [51,73,74]
CDSS Evalutions		
JAMIA	3	2 controlled trials [10,63]
		all designs [46,60]
Ann Int Med, Arch Fam Med, J. Am. Med. Assoc.	1 each	3 controlled trials [41,6,8], respectively
Guideline evaluations		
J. Am. Med. Assoc.	2	1 controlled trials [7]
		1 all designs [32]
Canadian Med Assn J	1	1 RCT [33]
Lancet	1	1 RCT [25]

*, +, †, §, ‡, Studies that appear to be related are marked with the same symbol and counted as one study.

References

[1] J.G. Anderson, S.J. (Eds.), Use and impact of computers in clinical medicine. Springer; New York, 1987.

[2] B. Kaplan, The medical computing 'lag': perceptions of barriers to the application of computers to medicine, Int. J. Technol. Assess. Health Care 3 (1) (1987) 123–136.

[3] B. Kaplan, The computer prescription: medical computing, public policy, and views of history, Sci. Technol. Human Values 20 (1) (1995) 5–38.

[4] A.G. Randolph, R.B. Haynes, J.C. Wyatt, D.J. Cook, G.H. Guyatt, Users' guides to the medical literature: XVIII. How to use an article evaluating the clinical impact of a computer-based clinical decision support system, J. Am. Med. Assoc. 282 (1) (1999) 67–74.

[5] E.A. Balas, S.A. Boren, Clinical trials and information interventions. In: E.S. Berner (Ed.), Clinical decision support systems: theory and practice. Springer, New York, 1999, pp. 199–216.

[6] E.A. Balas, S.M. Austin, J.A. Mitchell, B.G. Ewigman, K.D. Bopp, G.D. Brown, The clinical value of computerized information services, Arch. Fam. Med. 5 (1996) 271–278.

[7] D.A. Davis, M.A. Thomson, A.D. Oxman, B. Haynes, Changing physician performance: A systematic review of the effect of continuing medical education strategies, J. Am. Med. Assoc. 274 (9) (1995) 700–705.

[8] D.L. Hunt, R.B. Haynes, S.E. Hanna, K. Smith, Effects of computer-based clinical decision support systems on physician performance and patient outcomes: a systematic review, J. Am. Med. Assoc. 280 (15) (1998) 1339–1346.

[9] R.D. Zielstorff, M.J. Teich, M.D. Paterno, M.

Segal, P-CAPE: A high-level tool for entering and processing clinical practice guidelines, Proc. AMIA Symp. 1998; 478–82.

[10] S. Shea, W. DuMouchel, L. Bahamonde, A meta-analysis of 16 randomized controlled trials to evaluate computer-based clinical reminder systems for preventive care in the ambulatory setting, J. Am. Med. Inform. Assoc. 3 (6) (1996) 399–409.

[11] D.W. Bates, L.L. Lucian, D.J. Cullen, et al., Effect of computerized physician order entry and a team intervention on prevention of serious medication errors, J. Am. Med. Assoc. 280 (15) (1998) 1311–1316.

[12] D.W. Bates, J.M. Teich, J. Lee, D. Seger, et al., The impact of computerized physician order entry on medication error prevention, J. Am. Med. Inform. Assoc. 6 (4) (1999) 313–321.

[13] A.K. Jha, et al., Identifying adverse drug events, J. Am. Med. Inform. Assoc. 5 (2) (1998) 305–314.

[14] G.J. Kuperman, J.M. Teich, M.J. Tanasijevic, et al., Improving response to critical laboratory results with automation: results of a randomized controlled trial, J. Am. Med. Inform. Assoc. 6 (6) (1999) 512–522.

[15] M. Monane, D.M. Matthias, B.A. Nagle, M.A. Kelly, Improving prescribing patterns for the elderly through an online drug utilization review intervention: a system linking the physician, pharmacist, and computer, J. Am. Med. Assoc. 280 (1998) 1249–1252.

[16] R.A. Raschke, B. Gollihare, T.A. Wunderlich, A computer alert system to prevent injury from adverse drug events: development and evaluation in a community teaching hospital, J. Am. Med. Assoc. 280 (15) (1998) 1317–1320.

[17] H.R. Warner, Jr., S. Miller, K. Jennings, H. Lundsgaarde, Clinical event management using push technology—implementation and evaluation at two health care centers, Proc. AMIA Symp. 1998; 106-10.

[18] T.H. Lee, H.L. Cooper, Translating good advice into better practice, J. Am. Med. Assoc. 278 (23) (1997) 2108–2109.

[19] J.M. Overhage, W.M. Tierney, C.J. McDonald, Computer reminders to implement preventive care guidelines for hospitalized patients, Arch. Intern. Med. 156 (1996) 1551–1556.

[20] D.L. Schriger, L.J. Baraff, W.H. Rogers, S. Cretin, Implementation of clinical guidelines using a computer charting system: effect on the initial care of health care workers exposed to body fluids, J. Am. Med. Assoc. 278 (1997) 1585–1590.

[21] R.G. Woolf, R. Grol, A. Hutchinson, M. Eccles, J. Grimshaw, Potential benefits, limitations and harms of clinical guidelines, BMJ 318 (1999) 527–530.

[22] I.E. Buchan, et al., Local clinical guidelines: description and evaluation of a participative method for their development and implementation, Fam. Practice 13 (1) (1996) 28–34.

[23] D.C. Classen, Clinical decision support systems to improve clinical practice and quality of care, J. Am. Med. Assoc. 280 (15) (1998) 1360–1361.

[24] D. Cook, M. Giacomini, The trials and tribulations of clinical practice guidelines, J. Am. Med. Assoc. 281 (20) (1999) 1950–1951.

[25] J.M. Grimshaw, I.T. Russell, Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations, Lancet 342 (1993) 1317–1322.

[26] D.K. Litzelman, R.S. Dittus, M.E. Miller, W.M. Tierney, Requiring physicians to respond to computerized reminders improves their compliance with preventive care protocols, J. Gen. Intern. Med. 8 (1993) 311–317.

[27] D.F. Lobach, W.E. Hammond, Computerized decision support based on a clinical practice guideline improves compliance with care standards, Am. J. Med. 102 (1997) 89–98.

[28] M.M. Morgan, J. Goodson, G.O. Barnett, Long-term changes in compliance with clinical guidelines through computer-based reminders, Proc. AMIA Symp. 1998; 493–7

[29] P.J. Greco, J.M. Eisenberg, Changing physicians' practices, New Engl. J. Med. 329 (17) (1993) 1271–1274.

[30] C.L. McDonald, Protocol-based computer reminders, the quality of care, and the non-perfectability of man, New Engl. J. Med. 295 (24) (1975) 1351–1355.

[31] J. Bouaud, B. Séroussi, C.-C. Antoine, M. Gozy, D. Khayat, J.-F. Boisvieux, Hypertextual navigation operationalizaing generic clinical practice guidelines for patient-specific therapeutic decisions, Proc. AMIA Symp. 1998; 488–92.

[32] M.D. Cabana, et al., Why don't physicians follow clinical practice guidelines?: a framework for improvement, J. Am. Med. Assoc. 282 (1999) 1458–1465.

[33] D.A. Davis, A. Taylor-Vaisey, Translating guidelines into practice: A systematic review of theoretic concepts, practical experience and research evidence in the adoption of clinical practice guidelines, Can. Med. Assoc. J. 157 (4) (1997) 408–416.

[34] D.K. Litzelman, W.M. Tierney, Physicians' reasons for failing to comply with computerized preventive care guidelines, J. Gen. Intern. Med. 11 (8) (1996) 497–499.
[35] C.P. Friedman, A.S. Elstein, F.M. Wolf, et al., Enhancement of clinicians' diagnostic reasoning by computer-based consultation: a multisite study of 2 systems, J. Am. Med. Assoc. 282 (19) (1999) 1851–1855.
[36] J.P. Kassirer, A report card on computer-assisted diagnosis—the grade: C, New Engl. J. Med. 330 (25) (1994) 1824–1825.
[37] J. Ridderikhoff, E. van Herk, A diagnostic support system in general practice: is it feasible? Int. J. Med. Inform. 45 (1997) 133–143.
[38] F.M. Wolf, C.P. Friedman, A.S. Elstein, J.G. Miller, Changes in diagnostic decision-making after a computerized decision support consultation based on perceptions of need and helpfulness: a preliminary report, Proc. AMIA Symp. 1997; 263–7.
[39] V.L. Patel, Cognition and technology in health education research, Can. J. Public Health 87 (2) (1996) S63–S67.
[40] R.A. Garibaldi, Computers and the quality of Care—a clinicians' perspective. A computer-assisted management program for antibiotics and other antiinfective agents, New Engl. J. Med. 338 (4) (1998) 259–260.
[41] M.E. Johnston, K.B. Langton, B. Haynes, A. Mathieu, Effects of computer-based clinical decision support systems on clinical performance and patient outcome: A clinical appraisal of research, Ann. Intern. Med. 120 (1994) 135–142.
[42] J.S. Ash, P.H. Gorman, W.R. Hersh, Physician order entry in US hospitals, Proc. AMIA Symp. 1998; 235–39.
[43] E.S. Berner, R.S. Maisiak, G. Cobbs, O.D. Taunton, Effects of a decision support system on physicians' diagnostic performance, J. Am. Med. Inform. Assoc. 6 (5) (1999) 420–427.
[44] J.H. Carter, Design and implementation issues. In: Berner ES (Ed.), Clinical decision support systems: theory and practice. Springer, New York, 1999, pp. 169–197.
[45] H.A. Heathfield, J. Wyatt, Philosophies for the design and development of clinical decision-support systems, Meth. Inform. Med. 32 (1993) 1–8.
[46] R.A. Miller, Medical diagnostic decision support systems—past, present, and future: a threaded bibliography and brief commentary, J. Am. Med. Inform. Assoc. 1 (1994) 8–27.
[47] R.A. Nordyke, C.A. Kulikowski, An informatics-based chronic disease practice: case study of a 35-year computer-based longitudinal record system, J. Am. Med. Inform. Assoc. 5 (1) (1998) 88–103.
[48] C.M. Ruland, Improving patient outcomes by including patient preferences in nursing care. Proc. AMIA Symp. 1998:448-52; Decision support for patient preference-based care planning: effects on nursing care and patient outcomes. J. Am. Med. Inform. Assoc. 1999; 6 (4):304-312.
[49] J.G. Anderson, C.E. Aydin, S.J. Jay (Eds.), Evaluating health care information systems: methods and applications, Sage Publications, Thousand Oaks, CA, 1994.
[50] C.P. Friedman, J.C. Wyatt, Evaluation methods in medical informatics. Springer, New York, 1997.
[51] M. Berg, Rationalizing medical work: decision -support techniques and medical practices. MIT Press, Cambridge, 1997.
[52] B. Kaplan, R. Morelli, J. Goethe, Preliminary findings from an evaluation of the acceptability of an expert system in psychiatry, Extended Proc (CD-ROM) AMIA Symp (formerly SCAMC), 1997.
[53] T.W. Lauer, K. Joshi, T. Browdy, Use of the equity implementation model to review clinical system implementation efforts: A case report, JAMIA 7 (1) (2000) 91–102.
[54] C. Sicotte, J.L. Denis, P. Lehoux, The computer based patient record: a strategic issue in process innovation, J. Med. Syst. 22 (6) (1998) 431–443.
[55] C. Safran, P.C. Jones, D. Rind, B. Booker, K.N. Cytryn, V.L. Patel, Electronic communication and collaboration in a health care practice, Artif. Intell. Med. 12 (2) (1998) 137–151.
[56] C. Sicotte, R. Pineault, C. Tilquin, A.P. Contaridriopoulos, The diluting effect of medical work groups on feedback efficacy in changing physician's practice, J. Beh. Med. 19 (4) (1996) 367–383.
[57] C. Sicotte, J.L. Denis, P. Lehoux, F. Champagne, The computer-based patient record: challenges towards timeless and spaceless medical practice, J. Med. Syst. 22 (4) (1998) 237–256.
[58] B. Kaplan, Evaluating informatics applications—some alternative approaches: theory, social interactionism and call for methodological pluralism 64 (1) (2001) 15–37.
[59] E.S. Berner, Testing system accuracy. In: E.S. Berner (Ed.), Clinical decision support systems: theory and practice. Springer, New York, 1999, pp. 61–74.

[60] C. Nøhr, The evaluation of expert diagnostic systems—How to assess outcomes and quality parameters? Artif. Intell. Med. 6 (2) (1994) 123–135.
[61] R.P. van der Loo, Overview of published assessment and evaluation studies. In: E.M.S. van Gennip, J.L. Talmon (Eds.), Assessment and evaluation of information technologies, Studies in Health Technology and Informatics, vol. 17. Amsterdam IOS Press, 1995, pp. 261–282.
[62] R.N. Shiffman, C.A. Brandt, Y. Liaw, G.J. Corb, A design model for computer-based guideline implementation based on information management services, J. Am. Med. Inform. Assoc. 6 (2) (1999) 99–103.
[63] R.N. Shiffman, Y. Liaw, C.A. Brandt, G.J. Corb, Computer-based guideline implementation systems: a systematic review of functionality and effectiveness, J. Am. Med. Inform. Assoc. 6 (2) (1999) 104–114.
[64] D. Karlsson, C. Ekdahl, O. Wigertz, U. Forsum, A qualitative study of clinicians ways of using a decision-support system, Proc. AMIA Symp. 1997:268–72
[65] K. Clarke, et al., Methodology for evaluation of knowledge-based systems in medicine, Artif. Intell. Med. A 6 (2) (1994) 107–121.
[66] E.S. Berner, G.D. Webster, A.A. Shugerman, et al., Performance of four computer-based diagnostic systems, New Engl. J. Med. 330 (1994) 1792–1796.
[67] J. Talmon, et al., The VATAM guidelines, Int. J. Med. Inform. 56 (1999) 107–115.
[68] J. Concato, N. Shah, R.I. Horwitz, Randomized, controlled trials, observational studies, and the hierarchy of research designs, New Engl. J. Med. 42 (25) (2000) 1887–1892.
[69] S.J. Pocock, D.R. Elbourne, Randomized trials or observational tribulations?, New Engl. J. Med. 42 (25) (2000) 1907–1909.
[70] W.M. Tierney, J.M. Overhage, C.J. McDonald, An plea for controlled trials in medical informatics, J. Med. Inform. Assoc. 1 (4) (1994) 353–355.
[71] C.S. Gadd, P. Baskaran, D.F. Lobach, Identification of design features to enhance utilization and acceptance of systems for Internet-based decision support at the point of care, Proc. AMIA Symp. 1998; 91–5.
[72] L.D. Gamm, C.K. Barsukiewicz, K.H. Dansky, J.J. Vasey, J.E. Bisordi, P.C. Thompson, Pre- and post-control model research on end-users' satisfaction with an electronic medical record: preliminary results, Proc. AMIA Symp. 1998; 225–29.
[73] L.A. Moore, Reaction of medical personnel to a medical expert system for stroke. In: J.G. Anderson, C.E. Aydin, S.J. Jay (Eds.), Evaluating health care information systems: Approaches and applications. Sage, Thousand Oaks, CA, 1994, pp. 226–244.
[74] R.R. Weaver, Evaluating the Problem Knowledge Coupler: a case study. In: J.G. Anderson, C.E. Aydin, S.J. Jay (Eds.), Evaluating health care information systems: Approaches and applications. Sage, Thousand Oaks, CA, 1994, pp. 203–225.
[75] H.P. Lundsgaarde, Evaluating medical expert systems, Soc. Sci. Med. 24 (10) (1987) 805–819.
[76] J.C. Wyatt, D.J. Spiegelhalter, Evaluating medical expert systems: what to test and how?, Med. Inform. 15 (3) (1990) 205–217.
[77] W.W. Stead, R.B. Haynes, S. Fuller, C.P. Friedman, et al., Designing medical informatics research and library resource projects to increase what is learned, J. Am. Med. Inform. Assoc. 1 (1) (1994) 28–33.
[78] B. Kaplan, C.F. Flagle, Bibliography on Evaluation of Digital Imaging and Computer Based Medical Information Systems. Proposal to Develop an Evaluation Plan for the VistA Imaging System, The Department of Veterans Affairs, 1997–1998 (Project Co-Principle Investigator, with Charles D. Flagle, Johns Hopkins University. Subcontract between Birch & Davis Associates and Johns Hopkins University School of Hygiene and Public Health Subcontract 1523-001-S0-Hopkins), 1998.
[79] R. Kling, Social analyses of computing: theoretical perspectives in recent empirical research, ACM Comp. Surv. 12 (1980) 61–110.
[80] K. Lyytinen, Different perspectives on information systems: problems and solutions, ACM Comp. Surv. 19 (1987) 546.
[81] B.S. Mittman, X. Tonesk, P. Jacobson, Implementing clinical practice guidelines: social influence strategies and practitioner behavior change, QRB 18 (12) (1992) 413–442.
[82] E.R. Carson, Decision support systems in diabetes: a systems perspective, Comput. Methods Programs Biomed. 56 (2) (1998) 77–91.
[83] P.L. Miller, D.F. Sittig, The evaluation of clinical decision support systems: what is necessary versus what is interesting, Med. Inform. 15 (3) (1990) 185–190.
[84] G.C. Scott, L.A. Lenert, Extending contemporary decision support system designs to patient-oriented systems, Proc AMIA Symp. 1998, 376–80.
[85] D. Forsythe, B. Buchanan, Broadening our approach to evaluating medical expert systems, In:

P.D. Clayton (Ed.), Symposium on computer applications in medical care. McGraw-Hill, New York, 1991, pp. 8–12.
[86] H.A. Heathfield, I.E. Buchan, Current evaluations of information technology in health care are often inadequate, BMJ 313 (7063) (1996) 1008.
[87] R. Kling, W. Scacchi, The web of computing: computer technology as social organization, in: M.C. Yovits (Ed.), Advances in computers. Academic Press, New York, 1982, p. 290.
[88] A.W. Kushniruk, V.L. Patel, J.J. Cimino, in: D.R. Masys (Ed.), Usability testing in medical informatics: cognitive approaches to evaluation of information systems and user interfaces, Proc. AMIA Symp. 1997, pp. 218–22