Evaluating ICT Applications in Health Care

Studies from a Sociotechnical Perspective

Arjen Stoop

Evaluating ICT Applications in Health Care

Colofon

Financial support by Stichting BAZIS, PharmaPartners BV, Microbais BV and EuroNed for the publication of this thesis is gratefully acknowledged.
Chapters reprinted with permission of Schattauer GmbH, Germany (chapter 1) and Elsevier Ireland Ltd. (chapters 3 and 4).

ISBN 90-9019938-1

Printed by Universal Press, Veenendaal

Design and lay-out: Ineke Huibregtse
Cover: inspired by Joy Jones Kngwarreye, *Utopia*, 1996

Evaluating ICT Applications in Health Care: Studies from a Sociotechnical Perspective

Evaluatie van ICT toepassingen in de gezondheidszorg: studies vanuit een sociotechnisch perspectief

Proefschrift

ter verkrijging van de graad van doctor
aan de Erasmus Universiteit Rotterdam
op gezag van de rector magnificus Prof.dr. S.W.J. Lamberts
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden
op vrijdag 25 november 2005 om 13.30 uur

door

Arjen Peter Stoop

geboren te Echten (Friesland)

Promotiecommissie

Promotor:	Prof.dr. M. Berg
Overige leden:	Prof.dr.ir. A. Hasman Prof.dr. J. van der Lei Prof.dr. M. Trappenburg
Copromotor:	Dr. R. Bal

Dankwoord

Dit proefschrift is het resultaat van meerdere jaren werk bestaande uit het lezen en bediscussiëren van wetenschappelijke literatuur en andersoortige documenten, het observeren en analyseren van zorgpraktijken, het interviewen van actoren, het bijwonen van congressen en het schrijven (van vele versies!) van artikelen. Ondanks dat ik de meeste uren alleen heb doorgebracht was dit proefschrift niet tot stand gekomen zonder de hulp van vele anderen.

Om te beginnen wil ik de RITHM-groep bedanken. Binnen deze groep groeide mijn belangstelling snel voor onderzoek naar de mogelijkheden en het gebruik van informatietechnologie binnen de gezondheidszorg. Dat ik deel heb uitgemaakt van deze jonge, dynamische groep mensen met veel ruimte om te leren, heb ik als geweldig ervaren. Hierbij dank ik Sam Adams, Barbara Blank, Antoinette de Bont, Stans van Egmond, Mirjam Faber, Els Goorman, Bert Huisman, Yvonne Jansen, Femke Mastboom, Marleen de Mul, Habib Pirnejad, Irma van der Ploeg, Jolande Verhulst, Brit Ross Winthereik en Teun Zuiderent. Speciale dank gaat uit naar Jos Aarts die in de afrondende fase zelfs tot in de late uurtjes bereid was hulp te bieden, zowel inhoudelijk als praktisch.

From October 2002 to July 2004 I was a visiting Ph.D.-student at the Department of Health Services Research at the University of Copenhagen. I would like to thank Karsten Vrangbæk for making this possible and for welcoming me in a stimulating and friendly work environment. I also thank my other former colleagues for a very fruitful stay at the department: Lisa Dahlager, Tania Dræbel, Susanne Fray, Mogens Grønvold, Klaus Høyer, Morten Knudsen, Lene Koch, Allan Krasnik, Jette Michaelsen, Sniff Nexø, Anette Nielsen, Hans Okkels and Signild Vallgårda.

I also thank my co-authors and especially Heather Heathfield for their contribution to this thesis.

Alle data verkregen uit interviews, vragenlijsten en observaties heb ik verkregen door de belangeloze medewerking van vele enthousiaste respondenten. Het is helaas onmogelijk hen allen persoonlijk te bedanken voor de medewerking aan dit eindresultaat, maar ik ben hen zeer dankbaar.

Hiernaast dank ik meerdere mensen die weliswaar niet direct betrokken zijn geweest bij het schrijven van dit proefschrift, maar desalniettemin belangrijk zijn geweest. Allereerst dank ik AnneLoes van Staa voor haar inzet mij naar Rotterdam te halen en de prettige samenwerking op het gebied van onderwijs. Mijn latere kamergenoot, Radjesh Manna,

dank ik voor zijn humor en ontspannen manier van werken. Door hun positieve levensinstelling zijn mijn paranimfen Geertjan Straten en Jolande Verhulst zeer belangrijk geweest bij de verdediging en al hetgeen dat hierbij komt kijken. Tevens dank ik Ineke Huibregtse voor de snelle en deskundige hulp bij de vormgeving en opmaak van het proefschrift. Als laatsten in dit rijtje dank ik Patrick Geenjaar en mijn ouders voor al hun praktische hulp in de bepaald niet saaie periode waarin dit proefschrift tot stand is gekomen.

Als copromotor was Roland Bal, vooral in het laatste deel van het traject, zeer betrokken. Komend van een andere vakgebied was Roland in staat met een andere en verse blik te reflecteren op mijn stukken en nieuwe ideeën aan te dragen. Zijn hulp is voor mij zeer belangrijk geweest.

De belangrijkste reden voor mij om in 1999 naar het iBMG te komen was het aanbod van mijn promoter en begeleider Marc Berg om onderzoek te doen naar ICT toepassingen in de zorg. Aangezien ik Marc al meerdere jaren kende wist ik dat ik zijn aanbod niet kon weigeren. En ik ben hem nog steeds dankbaar: door zijn betrokkenheid, oneindige enthousiasme voor onderzoek en creatieve oplossingen voor ogenschijnlijk onoplosbare situaties, heb ik onze samenwerking als zeer prettig en constructief ervaren. Door niet te focussen op het eindresultaat, het proefschrift, maar op het uitvoeren van interessant onderzoek, is Marc gedurende het hele proces in staat geweest mij doelgericht en geïnspireerd te laten werken.

Ten slotte bedank ik vooral jou, Karen. Jij hebt mij tijdens het schrijven van dit proefschrift vele nieuwe en waardevolle perspectieven op het leven laten zien. Jouw doortastendheid en creativiteit zijn inspirerend. Het terugverhuizen naar Nederland - als gezin (!) - is cruciaal geweest voor het afronden van dit proefschrift. Terugkijkend en vooruitkijkend, ben ik er zeker van dat we nog vele andere mooie avonturen gaan beleven.

Amsterdam, september 2005

Contents

Introduction 1

Chapter 1 7
Integrating quantitative and qualitative methods in patient care information system evaluation: guidance for the organizational decision maker.

Chapter 2 21
Evaluation of Patient Care Information Systems: Theory and Practice.

Chapter 3 43
Using IT for Patient Education: Realizing surplus value?

Chapter 4 65
Theory and practice of waiting time data as a performance indicator in health care. A case study from The Netherlands.

Chapter 5 87
OZIS and the politics of safety: Using ICT to create a regionally accessible patient medication record.

Conclusion 105

Samenvatting 113

Curriculum Vitae 119

Introduction

This thesis is about evaluation of ICT applications in health care. The research described in this thesis started in the year 2000. In the research group RITHM (Research of Information Technology in Health care practice and Management) that was formed at that time, we discussed several aspects of the (potential) upcoming advantages of the use of ICT in health care compared to the current 'manual', 'paper-based' situation. Especially at that time, the advantages of the use of ICT in health care – both to clinicians and patients – seemed almost infinite. Electronic Patient Records (EPR's) promised to give access to patient data where ever and when ever the doctor liked [1]. Computerized Physician Order Entry (CPOE) systems would eliminate ambiguities caused by illegibility of handwritten orders and would make physicians cost-conscious by keeping prescribing practices in line with a hospital's established formulary [2]. ICT, in addition, would also offer the patient great opportunities like online accessibility to and control of medical data and electronic consultation of general practitioners by the Internet [3].

Although there had been a strong increase in the use of information systems in the last twenty years, the field of evaluation had received relatively little attention [4]. Many systems were not evaluated and those that were, were not (that) successful. In general, few health care institutions or vendors had been willing to spend money on the evaluation of information systems, despite low implementation success rates [5].

This situation has changed to a certain extent. The last few years there have been increasingly discussions about the role of evaluation studies for improving information systems, the use of evaluation methods and the need for evaluation frameworks and guidelines [6]. To put evaluation also on the political agenda, a group of experts has drafted 'The declaration of Innsbruck', which stressed the importance and the right conditions for evaluation studies [7]. Recently, because of new insights that show that information systems not only solve problems but also introduce new problems and dangers, authors have stressed the importance of evaluation studies to prevent hazards associated with ICT in health care [8]. Several scientific journals, finally, have become interested in publishing the results of evaluation studies.

This, however, does not mean that 'blueprints' of how to evaluate ICT applications in health care have arisen. The aim of this thesis is therefore to contribute to the above-mentioned discussions by using a sociotechnical perspective on the design, use and success and failure of ICT applications in health care. The sociotechnical approach is characterised by three starting points. First, (medical) work practices are conceptualized

as networks of people, materials, utilities, (professional and organisational) routines and so forth [9]. These elements should not be seen as discrete, well-circumscribed entities but as entities whose very nature is constituted in the network itself. The waiting room of a general practitioner as we know it, can only exist because of the way our health care is organized with the concept of appointments, efficiency and the 'rule' to give patients the opportunity to look for information while they are waiting. As a consequence, these entities cannot just be replaced without changing (elements of) the network. Replacing shelves with printed brochures by a computer-based information system, for example, changes the way that people (professionals or patients) look for information because different skills and knowledge are required.

Second, from a sociotechnical perspective medical work is characterised as ad hoc, pragmatic and fluid. These terms refer – like other complex types of work – to the fact that health care professionals constantly have to deal with contingencies that require ad hoc and pragmatic responses. This process is also called *managing a patient illness trajectory* and refers to all the work (doing investigations, diagnosing, intervening, monitoring) that has to be done to make patient care possible [10].

Third, a crucial part of medical work is perceived as 'invisible'. This invisibility refers to the work that – often done by nurses and assistants – has to been done to keep the trajectory going. It consists of tasks like making sure that the patient is physically and formally prepared for investigations or therapy. This work, however, is performed at the background and therefore often not recognized as important.

Applied to the evaluation of information systems, using a sociotechnical perspective means one stresses the importance of the interrelation between technology and its social and political environment to understand success and failure. Interrelation refers to how 'the technology' (the information system) and 'the social' (the users and the social and political context they are part of) interact. From a sociotechnical perspective, success and failure of information systems are the *outcome* of this interaction. The case studies in this thesis are illustrative for this statement.

In this thesis, ICT refers to applications that are meant to support (individuals and groups of) health care professionals in the primary care process as well as applications that are meant to inform and educate consumers and patients. Methods include observational studies, interviews, document analysis and questionnaires. Each of these methods is known for its strengths and weaknesses. Knowing these strengths and

weaknesses, one is able to capitalize on the strengths and – where relevant – compensate for the weaknesses.

This thesis consists of two parts. In the first part (chapters 1 and 2) I elaborate on theoretical standpoints regarding the current use of evaluation methods and an alternative way of *integrating* them. In addition, the difference between evaluation research in theory and evaluation research in practice is discussed by showing the complexity of evaluation in real-life settings. Moreover, I elaborate on the discussion about the required 'hardness' of evaluation data by describing the process of evaluation as a balancing act between limited resources, ever changing and multidimensional aims and the changing environment in which any project is situated. Because of this changing environment, I propose that evaluation has to be done *throughout* the implementation process (called formative evaluation), and not just *after* the implementation process (called summative evaluation) [11]. Formative evaluation makes it possible to adapt to changing circumstances during the implementation and consequently increases the chance of a successful implementation. The second part of this thesis (chapters 3, 4 and 5) consists of four case studies in which the methodological insights that have been outlined in the first part are applied and success and failure of the information systems is investigated. One of the often-mentioned outcomes of success and failure of information systems by using a sociotechnical perspective is the (mis)-match between the users' needs of (the functioning of) information systems according to the initiators/designers and those of the user itself. The inspiration to be aware of this potential mismatch – also in this thesis – is derived from the work of Diane Forsythe. In the 1990's she published several studies on this topic [12-14]. She showed how important it is to be aware of this potential mismatch in order to understand success and failure and to prevent – in case of a failure – to focus on the shortcomings of the user [15]. The four case studies consist of evaluation studies of different kind of ICT applications: two patient information systems, one webbased information system for patients and (individual) general practitioners and an information system for joint use by community pharmacists.

The following research questions are addressed:

1. In what way can evaluation studies benefit from the strengths and weaknesses of the different existing evaluation methods so they are better able to explain success and failure of information systems?

2. How can these insights be applied in real-life settings, what kind of possible constraints are observed and how can these constraints be dealt with?
3. Do computer-based information systems have the potential to realize surplus value compared to 'traditional' information systems used for education and information by patients and professionals and if not, how can this be explained?

These questions are addressed in the next five chapters.

Chapter one describes a general framework in which two important dimensions of evaluation are outlined: the domain of evaluation and the different phases of implementation. This chapter also shows how qualitative and quantitative methods can be *integrated* in such a way that one is able to capitalize on the strengths and weaknesses of each method.

Chapter two presents a six-step model on 'how to conduct an evaluation', using case studies from The Netherlands and the UK. The case studies clearly show how the theoretical insights in chapter one have to be related to the practical context in which evaluation takes place.

Chapter three concentrates on the potential benefits of ICT for patient education and shows it is crucial to be aware of all the tacit assumptions that accompany the design and use of information systems. This chapter analyses two patient information systems, both found to be unsuccessful because of a mismatch between the expectations of the designers and the needs of the users.

Chapter four is about the publication of information on waiting times on the Internet, meant to inform patients and general practitioners on choosing hospitals with the shortest waiting times. This chapter addresses the complexity of performance indicators and the need to do a thorough evaluation of information systems that are meant to create transparent information for the user.

Chapter five, finally, analyses a successful information systems for community pharmacists for exchanging information on patient medication data. This case study shows that alignment of the interests of the different stakeholders and political incentives are necessary for the information system to become *and* remain a success.

This thesis ends with a general conclusion.

References

1. Dick RS, Steen EB, ed. The Computer-Based Patient Record: An Essential Technology for Health Care. Washington, D.C.: National Academy Press, 1991.
2. Aarts JECM. Understanding implementation: a sociotechnical appraisal of the introduction of computerized physician order entry systems in Dutch and American hospitals. Doctoral thesis Erasmus University Rotterdam, 2005.
3. Kim M, Johnson KB. Personal health records: evaluation of functionality and utility. JAMIA 2002;9:171-80.
4. Kaplan B. Evaluating informatics applications - clinical decision support systems literature review. Int J Med Inform 2001;64:15-37.
5. Berg M. The Search for Synergy: Interrelating medical work and patient care information systems. Meth Inf Med 2003;42:337-44.
6. Kaplan B, Shaw NT. People, Organizational, and Social Issues: Evaluation as an exemplar. Yearbook of Medical Informatics 2002:91-102.
7. Ammenwerth E, Brender J, Nykanen P, Prokosch H-U, Rigby M, Talmon J. Visions and strategies to improve evaluation of health information systems: Reflections and lessons based on the HIS-EVAL workshop in Innsbruck. International Journal of Medical Informatics 2004;73(6):479-491.
8. Ammenwerth E, Shaw NT. Bad Health Informatics Can Kill - Is Evaluation the Answer? Meth Inf Med 2005;44:1-3.
9. Berg M. Patient care information systems and health care work: a sociotechnical approach. Intern J Med Inform 1999;55:87-101.
10. Strauss AL, Fagerhaugh S, Suczek B, Wiener C. Social organization of medical work. London: Transaction Publishers, 1997.
11. Friedman CP, Wyatt JC. Evaluation Methods in Medical Informatics. New York: Springer-Verlag, 1997.
12. Forsythe DE. Using ethnography to build a working system: rethinking basic design assumptions. Proc Annu Symp Comput Appl Med Care 1992:505-9.
13. Forsythe DE. Using ethnography in the design of an explanation System. Expert Syst Appl 1995;8:403-17.
14. Forsythe DE. New bottles, old wine: hidden cultural assumptions in a computerized explanation system for migraine sufferers. Med Anthropol Q 1996;10:551-74.
15. Forsythe DE. Blaming the user in medical informatics: The cultural nature of scientific practice. Knowledge and society 1992;9:95-111.

Chapter 1

Stoop AP, Berg M. Integrating quantitative and qualitative methods in patient care information system evaluation: guidance for the organizational decision maker. *Methods of Information in Medicine* 2003;42:458-62.

Abstract

With the current increase of Patient Care Information Systems (PCIS) in health care, the topic of evaluating such ICT applications becomes important. Yet the field of ICT evaluation is scattered: the types of questions that can be asked and methods that can be used seem infinite and badly demarcated. Different stakeholders, moreover, often have different priorities in evaluating ICT. The aim of this study is twofold. First, we describe two important dimensions of PCIS evaluation: the *domain* of evaluation and the *different phases* of the PCIS implementation. Second, we claim that, though Randomized Controlled Trials (RCTs) are often still seen as the standard approach, this type of design hardly generates relevant information for the organizational decision maker. The most important reason for this lack of relevance is that RCTs are based on controlled, laboratory conditions, and are well-suited for studying whether a particular intervention has a pre specified effect, but are not well-suited for investigating why and how a PCIS is being used, or not, and what the (often unplanned) effects and consequences are. Subsequently, our aim is to contribute to the discussion about the viability of qualitative versus quantitative methods in PCIS evaluation, by arguing for a specific integration of quantitative and qualitative research methods. The joint utilization of these methods, we claim, yields the richest results.

Keywords: Evaluation, Multimethod Approach, Randomized Controlled Trial, Patient Care Information System

Integrating Quantitative and Qualitative Methods in Patient Care Information System Evaluation: Guidance for the Organizational Decision Maker

Arjen Stoop, Marc Berg

1 Introduction

With the current increase of Patient Care Information Systems (PCISs)[1] in health care, the topic of *evaluating* such ICT applications becomes important. When it is claimed that a PCIS will enhance the quality and efficiency of care, for instance, it is important to assess whether this promise is indeed made true. Further, when we are implementing different PCISs, using different implementation methods, in different health care organizations, it is imperative that we learn from success and failure so that the same wheel is not re-invented each and every time. Yet the field of ICT evaluation is scattered; the types of questions that can be asked and methods that can be used seem infinite and badly demarcated. Different stakeholders, moreover, often have different priorities in evaluating ICT of which it is crucial to be aware of to avoid problems during evaluation [1]. Managers may want to know what the organizational impact is of a PCIS implementation process, and/or want to know whether their investment was economically worthwhile. Health care professionals might be primarily interested in patient outcomes, in workers' satisfaction, or in other quality-related indicators. Patients might be particularly interested in patient outcomes and patient satisfaction. An overall 'success' measure of information systems is rarely relevant [2].

Given these differences in questions and (interests of) stakeholders, the question how such an evaluation should be done has no easy answer.

[1] PCIS is a broader term than e.g. 'electronic patient record', but no sharp terminological distinctions are intended here. All these systems denote IT applications that handle information generated or used in the primary care process, and whose core users are doctors, nurses and other health care professionals (e.g. electronic patient records, order-entry systems, decision support systems).

In practice, many evaluation projects fail because the selected evaluation techniques can not properly answer the questions asked [3, 4]. Given the multitude of potential questions and priorities of different stakeholders, it is impossible to give a blueprint of how to proceed in designing an evaluation and selecting the proper method. However, this does not mean that it is impossible to provide guidance for making these methodological choices.

This study takes the perspective of the organizational decision-maker, who is confronted with the need to acquire, implement and/or manage a PCIS, and who wants to know what questions may be relevant, and how these may be addressed. We do not here, then, focus on evaluations with a primary scientific orientation, or, for example, on the comparative evaluations of classes of PCIS.

The aim of this chapter is twofold. We first describe two important dimensions of PCIS evaluation: the *domain* of evaluation and the *different phases* of the PCIS implementation. By domain we mean the different viewpoints an evaluation can take: it can focus on the *technical* performance of a system, for example, or on the impact of the system on *organizational* matters. The second dimension refers to the fact that evaluations can occur at different moments in the organization's dealing with the PCIS: before, during or after implementation. These categorizations are necessarily rough, and sometimes, actual implementation trajectories might be difficult to categorize as being either 'during' or 'after' implementation, for example. Equally, the domains that we distinguish should not be seen as neatly differentiated, mutually exclusive categories. Nevertheless, outlining these dimensions brings some order to the multitude of potential evaluation questions and can therefore help the organizational decision-maker that has to decide on evaluation[2].

Secondly, our aim is to contribute to the discussion about the viability of qualitative versus quantitative methods in PCIS evaluation, by arguing for a specific integration of quantitative and qualitative research methods. More often than not, these methodological discussions end in a polarized debate about the strong points of qualitative methods and the weaknesses of quantitative methods – or vice versa [5]. We recognize the deep paradigmatic differences that legitimate these debates, and that often complicate attempts to 'mix' methodological approaches [6, 7]. Yet we also agree with those authors that claim that combining the strengths

[2] Depending on the type of information technology that is evaluated (e.g. a PACS vs an EPR), some domains and effect measures are more suitable and relevant than others. See e.g. [21].

of different research methods can lead to richer research results and because of that can be an important step forward – both for the manager, who needs to make informed decisions about the future of the PCIS, as for the scientific understanding of PCIS implementation [8].

2 Domains and phases

When evaluating a PCIS, many different decisions have to be made: decisions about why to evaluate, what to evaluate, when to evaluate and how to evaluate are the most important ones. The question *why* to evaluate is important, because it should be clear at the outset what will be *done* with the evaluation's results in order to decide what type of results the evaluation should yield. When the aim of the introduction of a PCIS, for example, is to improve patient care, the evaluation should concentrate on parameters measuring patient outcomes and satisfaction before and after the introduction of the PCIS. This might seem very obvious, but the goals of evaluations are often not explicated, with ambiguous or meaningless results as a consequence [3]. In this chapter we will concentrate on the remaining three questions: *what* to evaluate, *when* to evaluate and – in the next paragraph - *how* to evaluate. First of all, different domains can be distinguished. By *domain* we refer to the different viewpoints that an evaluation can take: technical, professional, organizational, economic, ethical and legal. Though this list is not exhaustive, it does seem to cover all evaluation items that are relevant for the organizational decision-maker. Second, the *moment of evaluation* is crucial. Three phases can be distinguished: pre-implementation, during implementation and post-implementation. Within these phases of implementation of a PCIS, different evaluation questions can become relevant. Also, the overall aim of evaluation in these three phases is usually different. Evaluation during the pre-implementation phase, for example, can be done to test the feasibility of the intervention. It can also be carried out to decide whether or not to make a full evaluation later on [9]. In the implementation phase, evaluations often are concerned with providing feedback to help optimize the implementation process, which is called formative evaluation. In the post-implementation phase, evaluation is usually about the final outcomes or impacts of the intervention, and is called 'summative evaluation' [10].

When we combine these two dimensions, a table emerges in which each cell contains a distinctive set of questions (Table 1). In the pre-implementation phase, many questions are concerned with trying to find knockout arguments to decide in favor or against the system, and with

Table 1: Relevant evaluation questions

	Pre-implementation	**Implementation**	**Post-implementation**
Technical	- compatibility with other systems? - upgradable? - maintenance? - dataconsistency? - speed? - adaptability to changing requirements?	- possible to tailor information to specific needs of professionals? - downtime (frequency, duration)? - upgradable?	- how did the system perform on all selected features? - were there unexpected problems and were they solvable?
Professional	- what are the professionals' needs? - how much time does it take to learn the system/work with it? - does it make work more easy? - what are professionals' interests to work with the system? - user-friendliness? - content of information: understandable and complete?	- is it easy to use? - what are the benefits compared to the old situation? - (how) does it affect the content/effectiveness of work? - does it seem to improve patient outcomes (compliance, morbidity, mortality)?	- 'final' impacts on content/effectiveness of work (changing tasks, responsibilities, routines, less errors, time saving, less 'lost' records)? - improved data quality? - impact on patient outcomes? - overall satisfaction? - overall use?
Organizational	- is the organization ready? - are the different stakeholders ready? - is the objective of implementation clear? - does the investment 'fit' with other organizational strategies? - what kind of preparations/adjustments have to be made in advance?	- does the organization has to make adjustments (procedures, strategy, decision-making)? - are there unexpected negative effects?	- impacts on work processes and organization as a whole (communication patterns, responsibilities, decision-making procedures, interactions within/between professional groups)? - impacts on waiting lists, provided services, organizational strategy? - impacts on patient satisfaction?
Economic	- expected costs of buying, training, maintenance? - expected benefits (return on investment)? - reliable vendor?	- unexpected costs (maintenance, upgrading, training)? - are risks being managed?	- 'final' costs? - 'final' benefits?
Ethical	- data access, data security? - accountability for use of patient data? - possible effects of use of electronic patient data?	- how are patient data being used (by whom, for what purposes)? - who is responsible for use of patient data? - how are patients involved in the implementation?	- 'overall' effects of use of electronic patient data (e.g. decision-making, autonomy, doctor-patient communication)?
Legal	- expected registration (quality, presence) improvements/benefits?	- what role do electronic patient data play in legal matters (e.g. legal status)?	- consequences of use of electronic patient data? - legal (potential) (im)possibilities?

potential effects and expectations that need to be anticipated in thinking about an implementation strategy. In the implementation phase, questions are often concerned with the first consequences of real-time use and with tentative results. As mentioned, in this phase evaluation is geared towards optimizing the implementation process itself. In the post-implementation phase, it is geared toward accounting for and learning from decisions made. Here, attention primarily turns to overall outcome measurements.

3 Which methods to use?

What methods are most suitable for which types of questions? In general, qualitative research methods like interviews, observations and document analysis are optimally suited to understand a phenomenon 'from the points of view of the participants and in its particular social and institutional context' [11, p. 47]. They are capable of getting to the *what*, *why* and *how* of a social phenomenon: how users perceive and experience a system, for example, what the influence is of the social and organizational context on system use, or why an implementation strategy that worked in organization A does not work in organization B [11].

Quantitative research methods are most suitable for establishing the *size, extent* or *duration* of certain phenomena (*how much*), or to establish *that* a specific cause or intervention results in a prespecified effect. The 'how much' can be answered using different measurement techniques such as questionnaires, time studies or tracking of clinical outcomes. To establish a causal relationship, a broad range of more or less 'rigorous' designs are available, such as Randomized Controlled Trial (RCT), meta-analysis, cohort studies, case-control studies and observational studies.

There is no question that quantitative research methods have been the methods of choice in evaluating information systems. In the discussion between quantitative and qualitative research methods, it is important to separate the question of quantitative *measurement* per se from the question of study design[3]. After all, within one design, one can choose to use different methods. Generally, the Randomised Controlled Trial (RCT) is

[3] Words such as 'design', 'paradigm', 'methodology', 'method', 'technique' and 'measurement' are open to more than one interpretation (see e.g. [8]). In our paper, the word design refers to a description of the type of research (e.g. prescriptive, descriptive, longitudinal or cross-sectional) and the methods that are used (e.g. qualitative or quantitative).

seen to be the ultimate scientific design – the gold standard of evaluation [12, 13]. Yet this dominance has also been questioned for many years. First of all, social scientists have argued that 'scientificness' comes in many guises, of which hypothesis testing is only one. Building a theory explaining a specific social phenomenon, for example, is an equally scientific endeavour, which may or may not be amenable to quantification, or to 'testing' through a RCT. Second, the fact that conceptualizing (parts of) PCIS systems as 'the intervention' that a RCT focuses on supposes that it makes sense to isolate the functioning of this 'hardware' from the social processes that surround it and are integrated with its current functioning. This is a dubious assumption: in practice, it is often impossible to disentangle the 'effect' caused by the PCIS itself from the 'effect' caused by the changes in the workpractices induced by the PCIS implementation.

In addition to these general critiques, there are two additional reasons why RCTs may not be very useful evaluation tools for managers. First, a properly executed RCT is immensely labour intensive, and will give 'hard data' on (in the form of establishing relations between) a very limited set of pre-set parameters. It cannot answer the *why* or *how* questions that are often the most relevant when one wants to understand PCIS implementation, nor can it grasp all the unanticipated consequences that are often crucial to the fate of PCISs [13, 14]. RCT researchers themselves often stress that their designs are of limited 'real-world' use due to the artificial, laboratory circumstances (e.g. simulation patients, unexperienced subjects) in which the data are produced [15, 16].

In addition, conducting a RCT means that randomization has to be accomplished between two practices that are 'identical' except for the intervention that is being investigated (for example, randomly allocating patients to a clinician who uses a particular electronic patient record system and one who does not). However, generating these kind of 'objective' circumstances is impossible and unwanted in practice because of the peculiarities (routines, procedures, preferences) of professionals and departments within and between hospitals [6, 17]. More importantly, these peculiarities are exactly the reason why systems may fail in one situation, and may succeed in another – so our method should help us to know *more* about these issues, rather than *erase* them.

For all of these reasons, we argue that managerially oriented evaluations should emphasize designs that focus on qualitative research methods rather than RCTs, since qualitative methods are capable of generating insights that can explain (the effects of) those peculiarities. For example, grasping a phenomenon like user resistance (why does it

exist?, where does it exist of? and what are it effects?), can be done best by looking into those practices closely and using interviews and participant observations. Especially for the manager who needs to make decisions during the different phases of the implementation, this kind of information is crucial in order to optimize the implementation process.

However, less rigorous quantitative measurement techniques, compared to RCT's, can play an important role here. For example, in addition to qualitative results on user resistance (see above), for example, questionnaires can generate information on the amount of people that are unwilling or hesitant to use a system (on one moment or over a period of time). Ultimately, in our view, it is the *integration* of qualitative and quantitative methods that leads to the most valuable data, and, hence, to a deeper insight in the challenges and pitfalls of PCIS implementation projects. Combining qualitative and quantitative methods in information systems evaluation has been done - albeit infrequently - since the end of the '80s [14]. When this was done, the qualitative results were usually seen as the exploratory 'first steps' that could at best inform the 'real' research of generating hypotheses to be tested using experimental or statistical techniques. Yet more symmetrical integrations are now being proposed under the label of the 'multimethod' approach, where the different methods each produce their own data, without an implicit proposed hierarchy between them [18]. One of the reasons for the importance of 'multimethods research' is that the use of different methods is needed to capture the diverse and diffuse nature of information systems' effects. Another reason is to strengthen the robustness of research results through triangulation [19].

We agree with these insights, but we would like to argue that integration gains the richest results when the data from one method are used as *input* for the other. By using data as input we mean taking results from one method as a starting point for research of the other method. By doing this, it is possible to capitalize on the strong points of each method in order to gain more valid results and, consequently, strengthen the overall results. Qualitative research often is a prerequisite for quantitative research, because qualitative methods are best in identifying and selecting research topics for investigation. Quantitative research can, after that, be used to quantify these topics [20]. Subsequently, interpreting the results from quantitative research requires qualitative methods. To conclude whether the results can be regarded as 'bad' or 'good', for example, or to understand fluctuations or apparent contradiction in measured scores, qualitative interpretation is required to make sense of the obtained numbers. In our view, integrating quantitative and qualitative research

methods in this way requires that these quantitative designs should be quite 'modest' in nature: before-after designs are useful, but more strictly experimental designs overshoot their aim because these methods have severe limitations in the field of PCIS evaluation, as mentioned above.

We will illustrate how this integration can be done by using some of the questions from Table 1. We will focus on two domains that traditionally have been typically addressed with quantitative methods: by selecting 'hard cases' for our example, we hope to convince the reader that these insights are valid for the more 'soft' domains in our table as well.

From the *technical domain* for example, attention goes to software performance and hardware performance. From this focus it is interesting to know how stable the system is: how often and how long is it down, for example, and under what circumstances? After testing this during the pre-implementation phase under 'laboratory' circumstances, the real measurement can only be done during the implementation phase because the system has to be in real use (regarding frequency, duration and amount of usage, combination with other informations systems etc.). Since it is unclear at that moment to what extent (e.g. all functionalities or some functionalities) the system can go 'down' and under which circumstances, and since it is also hard to fully predict which performance issues matter most in the ongoing work of health care professionals, it is necessary to identify these issues qualitatively. After that, one can quantitatively measure frequency and duration of these performance issues. Subsequently, interpreting these results cannot be done without paying attention to the *consequences* of the performance (problems) on daily care. For example, do professionals mind that they have to create 'work-arounds' now and then? How much of a problem is it to have to wait a few seconds for a next screen? What are the key 'interrupting' performance issues for health care professionals? In the end, qualitatively interpreting the quantitative outcomes is the only way to generate the information that a manager needs to make an informed decision about the (required) technical performance of the PCIS.

Also for the *economic domain*, the integration of qualitative and quantitative research methods is obligatory. Although 'measuring costs and benefits' seems to be a straightforward, quantitative endeavour, it has become clear that what exactly *counts* as 'costs' or 'benefits', and how these should be valued, is not straightforward at all [21-23]. What counts as 'benefits', for example, can change over time because of changes in the organization and its context, of which IT is only a (small) part. Also, in practice many (e.g. strategic) benefits obtained from the introduction of new information systems appear to be unplanned [24]. Measuring the

costs of buying a PCIS, training personnel, maintenance etc., without looking at unplanned organizational changes like improved doctor-nurse communication, reduced waiting times or changes in tasks and responsibilities of staff produces incomplete results [25, 26]. The former mentioned costs (buying the PCIS, training of personnel, maintenance) can be identified and measured more or less precisely in the pre-implementation phase, whereas the latter mentioned costs can only be measured in the implementation and post-implementation phase. Identifying these latter costs and benefits, to prevent concentrating only on pre-specified indicators, requires qualitative methods. So even within the economical domain, qualitative research methods are required to identify (often less visible) consequences of implementing a PCIS. After identifying these, quantitative research methods can be used to measure costs like reduction of administrative staff, whereas qualitative methods can be used to analyze changes in doctor-nurse communication. Finally, interpreting these overall results, requires a qualitative as well as a quantitative approach: though the identified costs are important, organizational changes (e.g. changes in tasks or responsibilities) that sometimes hardly can be measured in terms of costs, can be perceived as more, less or equally important.

4 Conclusion

In our study we concluded that, though RCTs are often still seen as the standard approach for PCIS evaluation, this type of design is unsuitable for the organizational decision maker. The most important reason is that RCTs are based on controlled, laboratory conditions, and are well-suited for studying whether a particular cause or intervention has a pre specified effect, but are not well-suited for investigating why and how a PCIS is being used, or not, and what the (often unexpected) effects and consequences are. Since many PCIS fail and exactly this kind of information almost always is lacking, RCTs are hardly relevant from the managerial point of view.

In addition, we argued that in evaluating ICT applications two dimensions of PCIS evaluation are crucial: the *domain* of evaluation and the different *phases* of PCIS implementation. When we combine these two dimensions, a table emerges in which each cell contains a distinctive set of questions. By outlining these dimensions some order is given to the multitude of potential evaluation questions. Subsequently, the best way of answering these evaluation questions *and* interpreting results and consequences, we claim, is to integrate qualitative and quantitative research

methods in a specific way. By using results from one method as input for the other, richer results can be yielded compared to using these methods separately.

References

1. Talmon J, Enning J, Castãneda G, Eurlings F, Hoyer D, Nykänen P, et al. The VATAM guidelines. International Journal of Medical Informatics 1999;56:107-15.
2. Southon G. IT, Change and evaluation: an overview of the role of evaluation in health services. International Journal of Medical Informatics 1999;56:125-33.
3. Heathfield HA, Buchan IA. Current evaluations of information technology in health care are often inadequate. British Medical Journal 1996;313(19 October):1008.
4. Greatbatch D, Murphy E, Dingwall R. Evaluating medical information systems: ethnomethodological and interactionist approaches. Health Services Management Research 2001;14(3):181-91.
5. Savage J. Ethnography and health care. BMJ 2000;321:1400-2.
6. Heathfield H, Hudson P, Kay S, Mackay L, Marley T, Nicholson L, et al. Issues into multi-disciplinary assessment of healthcare information systems. Information Technology & People 1999;12(3):253-75.
7. Jones MR. Mission impossible? Pluralism and multiparadigm IS research. Information Systems Review 2000;1(1):217-32.
8. Mingers J. Combining IS Research Methods: Towards a Pluralist Methodology. Information Systems Research 2001;12:240-259.
9. Øvretveit J. Evaluating Health Interventions. Buckingham: Open University Press; 1998.
10. Friedman CP, Wyatt JC. Evaluation Methods in Medical Informatics. New York: Springer-Verlag; 1997.
11. Kaplan B, Maxwell J. Qualitative Research Methods for Evaluating Computer Information Systems. In: Anderson J, Aydin C, Jay S, editors. Evaluating Health Care Information Systems. California: SAGE Publications; 1994. p. 45-69.
12. Van der Loo R. Overview of published assessment and evaluation studies. In: Van Gennip E, Talmon J, editors. Assessment and evaluation of information technologies. Amsterdam: IOS Press; 1995. p. 261-83.
13. Kaplan B. Evaluating Informatics Applications: Review of the Clinical Decision Support Systems Evaluation Literature. International Journal of Medical Informatics 2001;"in press".

14. Kaplan B, Duchon D. Combining Qualitative and Quantitative Methods in Information Systems Research: A Case Study. MIS Quarterly 1988(December):571-86.
15. Friedman CP, Elstein AE, Wolf FM, Murphy GC, Franz TM, Heckerling PS, et al. Enhancement of Clinicianss Diagnostic Reasoning by Computer-Based Consultation: A Multisite Study of 2 Systems. Journal of the American Medical Association 1999;282(19).
16. Kuperman GJ, Teich JM, Tanasijevic MJ, Luf NM, Rittenberg E, Jha A, et al. Improving Response to Critical Laboratory Results with Automation: Results of a Randomized Controlled Trial. Journal of the American Medical Informatics Association 1999;6:512-22.
17. Berwick DM. Developing and Testing Changes in Delivery of Care. Annals of Internal Medicine 1998;128:651-56.
18. Anderson J, Aydin C, Jay S, editors. Evaluating Health Care Information Systems. California: SAGE Publications; 1994.
19. Kaplan B. Addressing Organizational Issues into the Evaluation of Medical Systems. Journal of the American Medical Informatics Association 1996;4(2):94-101.
20. Pope C. Qualitative Research: Reaching the parts other methods cannot reach: an introduction to qualitative methods in health and health services research. British Medical Journal 1995;311:42-5.
21. Jørgensen T. Measuring Effects. In: Van Gennip E, Talmon J, editors. Assessment and Evaluation of Information Technologies in Medicine. Amsterdam: IOS Press; 1994. p. 99-111.
22. Serafeimidis V, Smithson S. Information systems evaluation in practice: a case study of organizational change. Journal of Information Technology 2000;15:93-105.
23. Heathfield HA, Peel V, Hudson P, Kay S, Mackay L, Marley T, et al. Evaluating Large Scale Health Information Systems: From Practice Towards Theory. AMIA 1997:116-20.
24. Farbey B, Land F, Targett D. The moving staircase: Problems of appraisal and evaluation in a turbulent environment. Information Technology and People 1999;12(3):238-52.
25. Berg M. Patient care information systems and health care work: a sociotechnical approach. International Journal of Medical Informatics 1999;55:87-101.
26. Hebert M. Professional and Organizational Impact of Using Patient Care Information Systems. IMIA/Medinfo 1998:849-53.

Chapter 2

Adapted from:
Stoop AP, Heathfield H, De Mul M, Berg M. Evaluation of Patient Care Information Systems: Theory and Practice. In: Berg M, editor. *Health Information Management: Integrating Information and Communication Technology in Health Care Work.* London: Routledge, 2004.

Evaluation of Patient Care Information Systems. Theory and Practice

Arjen Stoop, Heather Heathfield, Marleen de Mul, Marc Berg

1 Introduction

Evaluation of patient care information systems (PCISs) has become increasingly important. Handling the increased complexity of health care processes, many argue, is impossible without the use of PCISs such as electronic patient records, patient data management systems, physician order entry systems and decision support systems. Many benefits are claimed: such systems will 'enhance the quality and efficiency of care', 'improve decision making of doctors', 'help reduce medical errors' and so forth [1-3]. Yet such claims can only be validated through evaluation of the performances and the effects of (using) these systems. In addition, the introduction and maintenance of PCISs consumes large amounts of resources and implementation failure is potentially a traumatic event for an organization. Decision-makers and those who are responsible for the procurement or development of IT are expected to demonstrate that resources spent on IT provide benefits in clinical outcomes, cost savings, and/or to the health care process. Furthermore, evaluation is increasingly important since there is a need to understand the effects of PCIS on the social, professional and organizational context in which they are used.

Those responsible for the design and implementation of evaluation studies, in the meantime, are faced with a bewildering and often conflicting array of choices and dilemmas concerning evaluation criteria, study designs, data collection methods and analysis techniques. In this chapter, we discuss some of the choices and dilemmas in PCIS evaluation, and provide the reader with practical guidance by presenting a six-step model. While evaluation can be aimed at many audiences, we here focus primarily on PCIS evaluation aimed at informing professionals and organization management in their decision-making about the development, implementation and use of the PCIS.

2 Formative evaluation of patient care information systems

In the field of evaluation of PCIS, two kinds of evaluation are distinguished: formative and summative. Formative evaluations focus on the continuous improvement of the system. Research is done throughout the lifecycle of a system, aimed at facilitating the organizational learning that is imperative for successful system design and implementation. Summative evaluations, on the other hand, are done to account *post hoc* for the promised or expected benefits, such as financial savings or the effectiveness of information systems in terms of clinical outcomes. Where formative evaluations are internally oriented, providing insights for the work groups, the project and steering group alike, summative evaluations are externally oriented, towards those paying or politically responsible for the PCIS [4].

Partly due to the historical entwinement with the RCT as 'golden standard', the PCIS evaluation literature still sharply divides 'objectivist' and 'subjectivist' approaches to PCIS evaluation (whether summative or formative) [5]. The objectivist position starts with the assumption that the merits and worth of an information system can and should be quantifiable. 'To measure is to know', it is often stated, and observations should be as objective as possible. In this way, it is argued, two observations of the same phenomenon will yield the same results, without being affected by how the resource under study functions or is used. Much attention, therefore, is given to the avoidance of measurement error, and the prevention of biased observation.

On the other side of the spectrum we find the subjectivist position, which argues that the impacts of health care ICT are often social and organizational in nature. In addition, even when one focuses on primarily medical or financial targets, social and organizational issues are co-responsible for the results. Therefore, some authors argue that proper health care ICT evaluation (summative and formative) requires the use of qualitative methods instead of quantitative methods. From this position, the results obtained from observation and interviews are context-dependent, and since the researcher is part of this context, 'objective data' are an illusion *in principle*. What is important from this position is to understand and document the different opinions that individuals and groups hold on issues, and the social and organizational processes that lead to a certain effect – such as the 'success' or 'failure' of a PCIS (see e.g. case study step 5 below) [6].

This perspective on using qualitative research methods is also present in formative PCIS evaluation, in which factors like limited time and resources, logistics, contradicting cultural, social and political forces between different stakeholders and so forth are considered to be research data [7]. For summative purposes, purely quantitative studies may be very useful. In the case of user resistance, for example, it is important to know how many users experience resistance and what the consequences of not using a system are for clinical outcomes. Also for summative purposes, however, it may be important to use qualitative methods as to know what the exact nature of the resistance is and what the impact of the user resistance is on the way professionals work within the organization. We claim therefore that for generating a balanced judgement, the use of quantitative *and* qualitative methods is required [8].

3 Steps in designing an evaluation

In this paragraph we will describe the design of an evaluation as following several, successive steps (as shown in Figure 1). We will focus primarily, but not exclusively, on formative evaluation. In reality, the individual steps are often difficult to separate, and usually impossible to manage as a simple and steady sequence. The process of designing an evaluation is one of adjustment and compromise, where choices concerning study design, evaluation criteria and data collection methods must be offset against the constraints of conflicting stakeholders' aims, deadlines, resources available, the intrusiveness of the evaluation and ethical considerations. The design of an evaluation is as much a social and political process as that of IT procurement and implementation itself. By using case material (from the UK and The Netherlands), we will show that taking these steps can be quite complicated. Every next step is determined by the situation at hand. Though the cases differ in technology, set up and success, both can be seen as a realistic example of the implementation of a PCIS in health care.

The more general 'Issues' and 'Lessons learned' at the end of each step are based on material from these two sites as well as other case material.

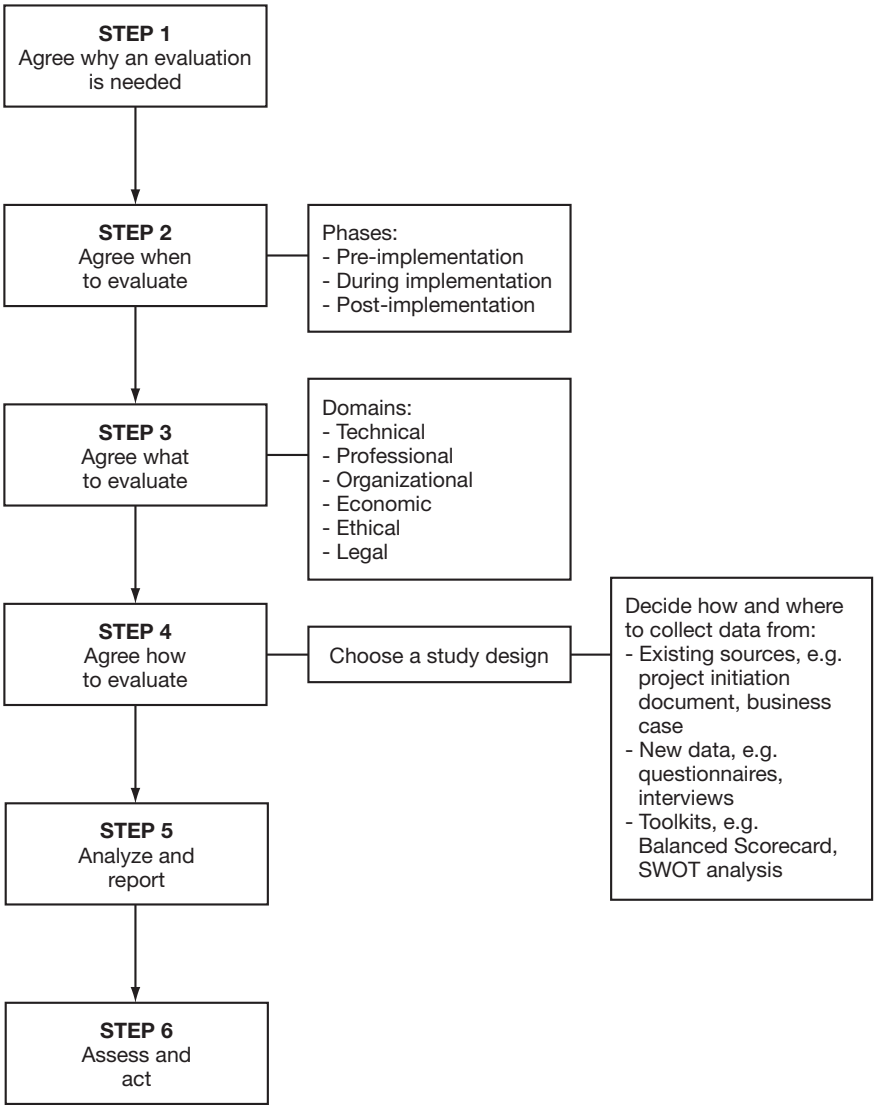

Figure 1: Steps to take for conducting an evaluation. (Source: Modified from [9])

Case study

The first case is based upon a study in which the UK NHS Executive, through its Information Management Group, initiated two programs of work which sought to improve patient records using information technology: the Electronic Patient Record (EPR) project and The Clinical Workstation (ICWS) project. The EPR

project was a three-year research and development program to assist clinicians in acute hospitals to provide better care to patients through the use of electronic patient record systems. The ICWS project was a two-year project concerned with providing interface facilities to an EPR. A number of sites with established Hospital Information Systems and sites in the early stages of procurement and implementation but with an organizational commitment to an EPR were selected. External evaluation was seen as essential to both the EPR and ICWS projects and the evaluation covered all five sites. Several stakeholder groups, at both national and local levels, were interested in the results of this evaluation. A multi-disciplinary team was assigned to evaluate the project, including health economists, social scientists, computer scientists and health services management specialists [10].

The second case refers to the late 1990s, when the Intensive Care Departments of a large Medical Centre in The Netherlands agreed to invest in information technology. The purchase of a Patient Data Management System (PDMS) was seen as an important first step towards an electronic patient record for Intensive Care, as it was a system to be used by both doctors and nurses. Like an EPR, a PDMS is an IT application that makes it possible to enter patient data. More importantly, the PDMS is explicitly focused on automatically storing and retrieving data from the heart monitor, the respirator and other electronic devices used in the Intensive Care Units. Moreover, the PDMS can also collect data from other data sources in the hospital, for example the laboratory results from the HIS. The system also makes it easier for the medical staff to order medication, because the system can calculate the right doses based on patient characteristics (age, weight). In addition, the system is able to make all sorts of automatically generated calculations, overviews and reminders, such as the fluid balance and medication overviews, and medication reminders. The system replaced an important part of the Intensive Care paper medical record, the bed chart [11].

Step 1 Agree why an evaluation is needed

This first step is crucial, because it does not only give direction to the goal of the evaluation and the questions that have to be answered, but, as a consequence, also to the required methods. If one wants to do an evaluation to compare a situation before and after the implementation

of a PCIS with the option to return to the old situation when the system does not function well enough, it is important to do a before–after study. This means that one first has to do a measurement of the situation before the implementation (baseline measurement), as to be able to make a good comparison between the situation with and without the system. A summative evaluation should be considered in these situations, as in this case the performance of the new system in use is to be compared to the situation where the system was not used. Formative methods can also be used in this case in order to facilitate better implementation of the system, but should be seen as complementary to the summative evaluation. If the implementation of a system is meant to replace work practices without the option of returning to the old situation without the system, for example, the formative evaluation of the system should focus on improving the way the system functions. In this case, a longitudinal design, throughout the whole implementation process, is appropriate.

Case study

In the EPR case study there was considerable difficulty in getting the stakeholders and commissioners of the evaluation to agree on the very reason for the evaluation. Many did not perceive an evaluation's role in promoting successful system implementation, for example. At last the team agreed on questions such as: What is the impact of the technology on clinical management regarding individual patient care, management of services and resource management? What is the experience of living and working at the implementation sites? During the study, however, it appeared that these questions were still far too general to answer directly, and that views on the concrete purpose of the evaluation were still diverging.

In the PDMS case the complexity of the question 'why evaluate' was caused by the fact that the evaluation had to serve different goals. On the one hand, the evaluation of the pilot had to deliver 'objective results' because the Board of Directors was about to decide about further investments in the rollout of the system to other ICUs. The project team therefore had to show the Board of Directors to what extent the goal of the pilot was met (the technical implementation of a PDMS with specific features and functionality). On the other hand, the members of the project team (especially the doctors and nurses) were interested in the impact of the system on work practices and user satisfaction. As a result, the evaluation of the PDMS pilot consisted of different parts, each

addressing different evaluation questions, requiring different methods and time scopes. Tying these different methods and time frames together proved difficult.

In this step, different issues come to the fore and several challenges have to be faced:

Issues
- Stakeholders may disagree on why an evaluation is needed, and it is not common that some question if evaluation is necessary at all (i.e., whether doing an evaluation is not a waste of already scarce resources).
- Different stakeholders may wish to evaluate for different reasons.

Challenges
- To make sure that all relevant stakeholders fully understand the role and importance of evaluation.
- The project team responsible for conducting the evaluation needs to have enough knowledge about conducting an evaluation.
- To be clear about how the evaluation results will be used, whether this is to inform and change the future direction of the project or to simply report and justify expenditure.
- To make the different stakeholders' perspectives and agendas explicit in the evaluation.

Step 2 Agree when to evaluate

The question when an evaluation is appropriate depends on different matters. First, it depends on the aim of the evaluation. Formative evaluations are geared towards process indicators and/or preliminary outcome measurements and are performed before and during the implementation of an information system. Summative evaluations are geared towards outcome parameters and are often performed before and after the implementation process to be able to compare both situations. Typical summative evaluations are, for example, cost-effectiveness or cost–benefit analyses.

Case study
In the UK EPR project, the evaluation was only commissioned after the start of the implementation, despite the fact that the commissioners of the work required the evaluation to demonstrate benefits in the post implementation as compared to the baseline situation. Therefore it was not possible to conduct a before–after study. In addition, due to the problem of defining and agreeing on

the scope of the evaluation, and the lack of initial access to the sites, the timescales for starting the evaluation slipped considerably.

The project team of the PDMS started to think about conducting the evaluation during the pilot implementation. Here also, when the decision to evaluate was made it was too late to do a before-evaluation of the 'old situation' (without the system). Consequently, topics like possible reduction of medical errors and completeness of medical records could not be evaluated. The project team decided therefore to evaluate user satisfaction during and after the implementation.

Issues

- Stakeholders often do not consider evaluation until late in a project which prevents a before–after study design.
- Evaluating early in the project life cycle can highlight problems and barriers, enabling them to be solved and the project to progress.

Challenges

- Start evaluation planning as early in the project life cycle as possible, preferably before project initiation.
- Getting 'evaluation' on the agenda this early in the project often requires an awareness of the organizational challenges of PCIS implementation that might not – yet – be present at the project and steering group level.

Step 3 Agree what to evaluate

Evaluations can be focused on different topics, which we will call domains. By domain we mean the different viewpoints an evaluation can take: it can focus on the technical performance of the system, on the impact of the system on professional or organizational matters, or on economic, ethical and legal features [8]. Looking at an EPR, for example, relevant topics from the technical domain are whether the system is compatible with other systems (like the hospital information system), whether it is easily upgradeable, whether it is easily maintained and whether the system has a fast response time. Relevant topics from the professional domain are whether the system meets professionals' needs, whether it is user-friendly, whether it makes the work of professionals easier. From the organizational domain it is relevant to know whether the organization is ready for the implementation of a new system, whether the objective of the implementation is clear and whether it fits

with the organizational strategy. From the economic domain it is interesting to know what the costs are of buying the system, training of personnel, maintenance and so forth. From an ethical point of view it is interesting to know what the effects of electronic patient data are on for example autonomy of the patient or the doctor–patient communication. From the legal domain, finally, topics like the legal status of electronic patient data are crucial.

When one simultaneously takes the domain and the phase of implementation in consideration (before, during or after implementation), a whole range of potentially relevant evaluation questions emerges (see chapter 1). After all, different kinds of questions are relevant during the different phases of the implementation trajectory. The importance of the different questions is dependent on the type of technology (e.g. an Electronic Patient Record versus a Patient Information System) and the focus of the evaluation [12]. Paying attention to all these questions offers the possibility to do a comprehensive evaluation, which means one can judge the information system on all relevant aspects. Since one can ask so many questions, the relevance of the evaluation questions is dependent on the perspective that one takes. There can be as many perspectives as there are stakeholders in an evaluation project. For example, evaluation questions regarding the technical aspects of maintenance, compatibility with other systems and possibilities to upgrade are relevant for the IT department, but generally less of a priority for professionals (doctors, nurses) or the manager that is accountable for the cost-effectiveness of an information system. Patients, for example, are interested in patient outcomes, but probably less interested in the expected costs of buying or the reliability of the vendor. Independent of the reason for the evaluation, the aim of evaluation is to find answers to relevant evaluation questions. This sounds incredibly obvious, yet *defining* what a relevant research question *is* in a given situation often appears to be far from easy. Since it is often less than clear *why*, from the perspective of the (actors within the) organization, a PCIS project is embarked upon, it is also often difficult to establish what should be evaluated. Sometimes, this problem can even lead to not conducting the evaluation at all [13]. However, once this is clear, it becomes relatively easy: systems that are implemented to enhance the quality of care in terms of, e.g., effectiveness, efficiency or satisfaction, should be evaluated by questions that are also geared towards these parameters. An EPR that is primarily implemented to enhance the ability to retrieve data should be evaluated by, e.g., looking at the speed, reliability, completeness and user friendliness of retrieving

data. Conducting an evaluation on all these different domains in the different phases of the implementation is in principle impossible – unless resources are unlimited. In addition, doing a comprehensive evaluation requires much knowledge and skills of the evaluators. As a consequence, designing the evaluation implies selecting only the relevant domains and phases.

Case study
A problem that we saw above arises here in a different guise. In the EPR case study one of the evaluation questions looked at the impact of the EPR on work practices. However, the question was posed as a general question, and did not specifically address questions of interaction between the user and the EPR, duplication of information recording or clinical uses of the information held in the EPR.Therefore it was unclear what the exact focus of the evaluation was. To gain clear results from an evaluation, the topic one evaluates should be very specific. There was also a push from the clients to look at clinical outcomes, although these were unlikely to show up until much later in the system lifecycle.

In the PDMS case the project team did not do a baseline measurement, which made before–after measurements impossible and, consequently, limited the range of possible evaluation questions. Other questions could not be addressed because there was not sufficient knowledge on how to set up an appropriate evaluation. This was the case for economic questions like 'what is the result of the investment in terms of saved resources (staff, time and money)' which was 'promised' by the vendor of the system. Therefore, they decided to focus only on two aspects of the system during and after the implementation phase: doctors and nurses' satisfaction with the system and their opinions of the system's effects.

In this step, the following issues and challenges come to play:

Issues
- Conducting a comprehensive evaluation is generally impossible because resources (time, knowledge, money) are always limited.
- Some areas are more difficult to evaluate than others. In particular, the effects of the introduction of a PCIS on clinical outcomes or costs-benefits are hard to measure because these effects take much time to become apparent.

Challenges

- It should be clear from the beginning what the topic of the evaluation is.
- It is important, given the limited amount of resources, to jump upon opportunities to conduct parts of the evaluation (data already gathered, meetings already planned where focus interviews can be held, and so forth).
- The project team has to agree on (and back up) the relevance and feasibility of evaluation questions.
- Evaluation questions have to be specific enough as to generate meaningful data.

Step 4 Agree how to evaluate

Choose a study approach

We have already claimed how important it is to define evaluation questions that are relevant, specific and feasible. Subsequently, the content of the question determines which methods have to be applied. Crucial, in these choices, is the question how – scientifically speaking – 'hard' the data need to be. Though RCTs from the subjectivist position are seen as the gold standard generating the 'hardest' data possible, we have argued that this design is unsuitable in the field of evaluating PCIS [8]. In practice there has to be a balance between the wish to generate robust data and the reasonableness of the effort to generate these data. Experimental designs, for example, generate 'hard' data compared to non-experimental designs like case-control or before–after studies, but they also require much more effort.

Choosing the study approach, then, is essentially a balancing act between the resources available and the breadth and depth of evaluation information required. Evaluating whether the use of an electronic prescribing medication system leads to gain or loss of time and reduction of errors, for example, would ideally require a design in which doctors' prescribing behaviour is precisely measured and compared between the old and new situation. However, interviewing doctors and those that are involved in reading and acting upon the prescriptions probably gives data that are hard enough to conclude whether the system meets its objective. When we were asked to evaluate the impact of a patient information system in the waiting room of the general practitioner, for example, we were asked to consider the impact in terms of efficiency (a reduction of consultations) and patient satisfaction. A fully elaborated design would require a comparative, longitudinal study. Because of limited resources, we resorted to doing a quick scan of the practices

using the system, through brief participant observations in the waiting room and interviewing GPs, secretaries and patients. This gave us already much information – enough to give the feedback that the system would not meet its objectives (see chapter 3). Heavily leaning on 'less than ideal (in the sense of "optimally objective")' designs is not a weak point of formative evaluations. To the contrary: when the evaluation questions are clear, and the evaluators and the members of the project and steering group have a clear understanding of one another's needs and possibilities, a little information can be very powerful. Implementing a PCIS is a highly uncertain and unpredictable process, and being able to obtain rapid information from a broad array of relevant issues can be priceless [14]. More 'objective' data would not be able to be at hand when managers actually need it, nor would it be possible to paint the breadth of the issues at stake in the implementation when only focusing on a few parameters measured.

Determine the right data collection method
After having chosen the right design, the question is which methods are most suitable to answer the evaluation questions. If one, for example, wants to investigate user satisfaction, questionnaires often are suitable. One can also do observations or a combination of these two methods. If the aim of the system is to improve clinical outcomes, one can review clinical records and outcome measurements or ask health-care professionals and patients what their perception is. If the aim is to improve completeness, a comparative study with a sample of paper records and electronic records is appropriate. The choice of methods is also depending on the scope of the evaluation: questionnaires and interviews make it possible to cover many topics, whereas time studies (e.g. on the amount of time it takes to use an electronic prescribing system compared to a paper method of prescribing) generate data on few aspects of the implementation. Here again, of course, the issue of 'balancing the costs versus the outcomes' in the evaluation design is crucial.

Decide how to collect data
Data for evaluation studies may come from many sources, some of which exist already, some of which have to be collected from scratch. Existing sources of data are useful for identifying criteria against which to evaluate. Potentially interesting documents are: local, sub-national and national benchmarks on reference costs or quality and performance indicators, project initiation documents, local or national standards. When there are no (available) data one can decide to collect new data,

for example by means of questionnaires, interviews, observations, focus groups, workshops, etc. One can also use self-assessment tools, like balanced scorecards, Strengths Weaknesses Opportunities and Threats (SWOT) analysis and benchmarks, which allows stakeholders to identify the project criteria that are important to them, and encourage them to think about these issues in depth. The right size of the study group, finally, depends upon the scope of the system being developed. If one is evaluating a system that is meant to cover a large number of patients but only regarding minimal data items, the potential sample size is large but the potential impact of the system may be small given the fact that the data may be useful only in a limited number of clinical situations. If the system is meant to cover a small number of patients but containing in-depth data, the potential sample size is small but the potential impact on clinical decision-making is greater.

This step, then, is not primarily a question of scientific rigour, but primarily a matter of balancing between the often-limited resources of the site, and the evaluation questions that the steering or project group wants to be answered. The choice for a study design is dependent on the time, knowledge and financial possibilities of the evaluation site. Designs like control and intervention studies regarding the implementation of for example an EPR are not realistic (how to find sites that, except for the application, are really comparable?). Regarding the scope of the study, it is again important to balance between the often-limited resources and the multitude of possible effects of a system by focusing on those domains that – looking at the evaluation question – are seen as most relevant. Focusing in-depth on one or two effects is dangerous because many other – more important – aspects of the use of the information system might then be overlooked. In addition, it is important to know the relation between the different aspects of using the system. Some effects of the use of the information technology may be caused by each other, e.g. time gain and satisfaction of the user, but others may not be related, e.g. increased efficiency of work practices and the impact on patient outcomes.

Case study

In the UK EPR case, a number of different techniques were used to answer the evaluation questions. But because these techniques were not in any way adapted to each other there was no match between the study design and the evaluation questions. Hence, it was not feasible to conduct any kind of comparative study with similar non-EPR sites, as each site is unique and the resources available for the evaluation were insufficient to cover the additional control sites.

The project team responsible for the PDMS implementation made a questionnaire for nurses and doctors. This questionnaire was based on a validated questionnaire that was, however, not aimed at information technology, but at a totally different topic. As a consequence, the questions of the original questionnaire had to be rephrased, some questions were left out and some new questions were added. This new questionnaire was seen as suitable for the PDMS context. For the evaluation of the technical success of the implementation, a practical study design was found: the functional demands (which were found in an initiation document) were compared to the PDMS-in-use.

Issues
- Stakeholders often assume that a (randomized) controlled trial is the only way of evaluating a system.
- There may be a lack of understanding about how to match study designs and evaluation questions.
- Stakeholders might underestimate the skills and time needed to design questionnaires and conduct interviews and focus groups.

Challenges
- To make sure that the evaluation resources are appropriate to the size and type of project.
- The project and steering group have to be convinced that a pragmatic research design is the most optimal research design for formative evaluation projects.

Step 5 Analyse and report

The interpretation of the data can be complicated. First, even when one has a very clear and limited research question, the effects of the implementation of applications can be diverse and unexpected. Several elements (such as the culture of a hospital regarding innovation, the computer literacy of professionals, work satisfaction or the way professionals work together) can impact the implementation, and should therefore be included in the evaluation [15,16]. Yet even when this is done, it often remains a difficult task to ascribe the impact of the different elements to the outcomes measured [17]. Second, though from the data it may appear that the system has no impact, this does not mean that this is the case. For example, results may be contradictory or unclear which makes it impossible to draw valid conclusions. It may also be that though the parameters used in the evaluation do not show any impact, professionals

or patients do perceive a difference (on parameters not measured). Third, the object of evaluation often steadily changes. Because of staff or workflow changes, software modification, training sessions, etc., the information system may change and be in the end markedly different from what it was at the beginning of the study [18]. In addition, for many applications it is generally difficult to make statements about the success or failure of a system immediately after the implementation phase, because it often takes years for systems to have its maximum impact. It takes time for work processes to find an optimal equilibrium with the new system, for conceptual, design and implementation errors to be repaired, and so forth. This is called the 'evaluation paradox': although it is desirable to get evaluation results as soon as possible so that one can decide on following steps (e.g. adjustments or aborting the system), it is often impossible to generate 'final' results within short time [5].

Case study

The EPR case study illustrated the problem of not presenting results as accessible for the client or the commission. Instead of a clear and practical guidance in the form of recommendations, a large report was delivered with different kinds of results – ranging from mainly descriptive material about the way health care professionals work (many hundreds of pages) to short technical reviews, e.g. on the robustness of the system. As we already said, the results were impossible to triangulate. Subsequently, it was left to the client to search through and identify the salient points (which of course was unrealistic). Furthermore, the report was delivered after major decisions about future EPR initiatives were made.

The results of the PDMS case illustrated the differences in perception of success of the system between and even within groups of stakeholders, for example regarding time investment in the system. Though the system made several manual activities superfluous (like making and checking the 24 hours-list, or copying the medication list and the nursing orders list to every new bed chart), these activities were normally conducted during the nightshift and the gained time remained therefore unnoticed for the dayshift. Also, because of readability and automatic storage of prescriptions, the use of the PDMS on one location (the ICU ward) saved time on another (the pharmacy connected to the ICU ward). Bringing such issues to light is often crucial to prevent mutual frustrations to fester.

Issues
- The implementation of information technology changes both work practices and the technology itself. The phenomenon to be investigated, then, is not stable in formative PCIS evaluation.
- The way in which evaluation data is presented can skew the interpretation.

Challenges
- It is important that evaluation results are presented in a clear and concise manner.
- Sometimes, different reports should be produced for different stakeholders, reflecting their perspectives and need for detail.
- Different results should, where possible, be linked to each other. For example, a qualitative result may explain a quantitative finding.

Step 6 Assess recommendations and decide on actions

An important part of evaluation is the consideration of the implications of the evaluation. By doing this in the right way, continuous improvement and development of a learning culture is stimulated. If we want to learn from evaluation studies, we not only have to publish negative results, but also act on them. This is not always easy, of course. Evaluation results might lead to the conclusion that, looking at the aims of the PCIS, it is best to abort the system. For example, alerting systems that appear to be unreliable or alert too often too fast will not be accepted by personnel or management of a hospital. What is needed, therefore, is a formal documented action plan that is agreed upon by all stakeholders and allocates responsibilities for improvement and identifies timescales. In addition, it is important that there is proper communication of the actions and that necessary adjustments are made in, for example, policies in order to make the actions possible. Finally, this step should be seen as establishing a – new – baseline that is crucial for the next steps in a possible next evaluation.

For the interest of the health care community, it is important that negative findings are also viewed as a basis for shared learning and action planning. Since many local – more or less identical – initiatives are undertaken, it is especially important that these experiences are documented and people are informed about systems and implementation trajectories that are successful and those that have failed [19]. Doing this requires a thorough analysis of the reasons for success or failure of PCIS.

Case study

In practice, however, negative results are often not published and positive results are not always acted upon. In the UK EPR study, clients were keen to see positive results and the sites that were studied did not wish their projects to be viewed in anything but the most positive light. This made it difficult to present specific examples, lending less weight to the findings.

In the PDMS case, the results of the survey were positive. The evaluation of changing work practices and user satisfaction resulted in several technical improvements of the system and in better communication to the users. However, despite the positive results of the evaluation, the Board of Directors decided not to prioritize the implementation of the PDMS on three other ICU wards, the operating rooms and the Emergency Departments. Faced with a small budget, the project team could only install the system on two wards.

Issues

- Different stakeholders may selectively focus on specific bits of information out of the overall context of the evaluation report, to illustrate their own points.
- Despite positive evaluation results, further steps might be frustrated.
- Often it is unclear who is responsible for taking action on the evaluation results.

Challenges

- Clear communication on the core of the evaluation results and appropriate action is crucial.
- Defining who is responsible for receiving and acting upon evaluation results should be done at the evaluation planning stage – not later once results appear.

4 Conclusion

In this chapter, we have provided an introduction to the theory and practice of evaluation. To this end, we have given the reader an appreciation of some of the major debates in this area. In addition, we have showed that conducting an evaluation is not simply about following several steps in a certain sequence. Designing and conducting an evaluation is a balancing act between identifying specific and feasible evaluation questions, utilizing the amount of resources available and specifying the sufficient

'objectivity' of the data. However, distinguishing the steps is useful since they constitute a framework through which to introduce and discuss the relevant issues. In addition, the steps are equally useful as a framework to guide one's own work. Yet the overriding requirement in the practice of designing and doing an evaluation is the concrete balancing act between limited resources, ever changing and multidimensional aims, and the changing environment within which any project takes place.

References

1. Dick RS, Steen EB, ed. The Computer-Based Patient Record: An Essential Technology for Health Care. Washington, D.C.: National Academy Press, 1991.
2. Bobb A, Gleason K, Husch M, Feinglass J, Yarnold PR, Noskin GA. The Epidemiology of Prescribing Errors: The Potential Impact of Computerized Prescriber Order Entry. Arch Intern Med 2004; 164(7):785-792.
3. Bates D, Pappius E, Kuperman G, Sittig D, Burstin H. Using information systems to measure and improve quality. International Journal of Medical Informatics 1999;53:115-24.
4. Beynon-Davies P, Lloyd-Williams M. When health information systems fail. Top Health Inf Manage 1999;20:66-79.
5. Friedman CP, Wyatt JC. Evaluation Methods in Medical Informatics. New York: Springer-Verlag, 1997.
6. Beynon-Davies P. Information systems 'failure': the case of the London Ambulance Service's Computer Aided Despatch project. European Journal of Information Systems 1995;4:171-84.
7. Berg M. Patient care information systems and health care work: a sociotechnical approach. Intern J Med Inform 1999;55:87-101.
8. Stoop AP, Berg M. Integrating quantitative and qualitative methods in patient care information system evaluation: guidance for the organizational decision maker. Meth Inf Med 2003;42:458-62.
9. Heathfield H, Clamp S, Felton D. Probe: Project Review and Objective Evaluation for Electronic Patient and Health Record Projects.: UK Institute of Health Informatics for the NHS Information Authority, 2001.
10. Heathfield H, Hudson P, Kay S, Mackay L, Marley T, Nicholson L, Peel V, Roberts R, Williams J. Issues into multi-disciplinary assessment of healthcare information systems. Information Technology & People 1999;12(3):253-75.

11. De Mul M, Berg M, Hazelzet JA. Clinical information systems: CareSuite from Picis. Journal of Critical Care 2004;19:208-14.
12. Jørgensen T. Measuring Effects. In: Van Gennip E, Talmon J, ed. Assessment and Evaluation of Information Technologies in Medicine. Amsterdam: IOS Press, 1994:99-111.
13. May C, Ellis NT. When protocols fail: technical evaluation, biomedical knowledge, and the social production of 'facts' about a telemedicine clinic. Social Science & Medicine 2001;53:989-1002.
14. Berwick DM. Developing and Testing Changes in Delivery of Care. Annals of Internal Medicine 1998;128:651-56.
15. Kaplan B, Shaw NT. People, Organizational, and Social Issues: Evaluation as an exemplar. Yearbook of Medical Informatics 2002: 91-102.
16. Kaplan B, Shaw NT. Future Directions in Evaluation Research: People, Organizational, and Social Issues. Meth Inf Med 2004;43: 215-31.
17. Ammenwerth E, Graber S, Herrmann G, Burkle T, Konig J. Evaluation of health information systems--problems and challenges. International Journal of Medical Informatics 2003;71(2-3):125-135.
18. Farbey B, Land F, Targett D. The moving staircase: Problems of appraisal and evaluation in a turbulent environment. Information Technology and People 1999;12(3):238-52.
19. Van't Riet A, Berg M, Hiddema F, Sol K. Meeting patients' needs with patient information systems: potential benefits of qualitative research methods. Intern J Med Inform 2001;64:1-14.

Chapter 3

Stoop AP, Van't Riet A, Berg M. Using IT for Patient Education: Realizing surplus value? *Patient Education and Counseling* 2004;54:187-195.

Abstract

Computer-based patient information systems are introduced to replace traditional forms of patient education like brochures, leaflets, videotapes and, to a certain extent, face-to-face communication. In this study, we claim that though computer-based patient information systems potentially have many advantages compared to traditional means, the surplus value of these systems is much harder to realize than often expected. By reporting on two computer-based patient information systems, both found to be unsuccessful, we will show that building computer-based patient information systems for patient education requires a thorough analysis of the advantages and limitations of IT compared to traditional forms of patient education. When this condition is fulfilled, however, these systems have the potential to improve health status and to be a valuable supplement to (rather than a substitute for) traditional means of patient education.

Keywords: Patient education; Computer-based patient information systems; Information technology

Using information technology for patient education: realizing surplus value?

Arjen Stoop, Annemarie van't Riet, Marc Berg

1 Introduction

Adequate patient information is important for the quality of care. It is one of the key indicators of patient centeredness and it improves the effectiveness and efficiency of care [1]. Until approximately 10 years ago, patient information was given face-to-face or through different means like paper-based flyers, brochures and videos. Since a few years, computer-based patient information systems (for teaching, decision support, information retrieval, etc.) have become more and more available. These new type of systems, several authors claim, have many advantages compared to the more traditional means used for patient education. The biggest advantage of these systems is that they are able to use more interactive ways of informing patients and are better able to tailor information to the individual needs of patients at less costs [2–6]. Also, by using advanced search techniques, computer-based patient information systems have the potential to store and retrieve large amounts of information. In addition, computer-based patient information systems offer patients new possibilities of getting in contact with other patients or experts by establishing virtual meeting groups or discussion lists. Though little is known about concrete effects of using these new type of systems, aspects such as patient decision-making and planning, patients' trust in the caregiver, compliance and patients' motivation seem to be improved [7–9].

One of the reasons that there is not much clear evidence of the potential benefits of the use of patient information systems is that few studies focus on the *evaluation* of computer-based patient education. In a review article on computer-based approaches to patient education Lewis found only 21 research-based reports between 1971 and 1998 that included evaluation findings [10]. In addition, results from evaluation studies of computer-based patient education show different results. Though some studies for example point at improved transfer of knowledge and increased patients' expertise, others state that computer-based education

does not provide additional significant gains compared to traditional forms such as face-to-face education or non-tailored patient information [10,11]. From the literature, it appears to be quite a challenging task to build information systems that meet patients' needs. Designers, as Diane Forsyth claims, often hold tacit assumptions about what patients want or need that appear problematic for users [12]. As a consequence, the utility of such systems is strongly diminished [13]. Explicating patients' needs, however, is not an easy task. To individualize patient education materials, designers must consider the unique needs of the target audience to include culture, age, race, gender as well as social issues and physical and psychological or cognitive disabilities [10,14].

In this chapter, we want to report on two evaluation projects. Though there are differences in the two projects regarding for example the type of information system and the potential user (see further), the goals of the two systems were comparable and the results are complementary. We will show that the failure of both patient information systems was caused by the fact that the designers had not made a thorough analysis of the advantages and limitations of IT compared to traditional means for patient education. As a consequence, they built computer-based patient information systems that did not offer surplus value to the available educational means. Though with hindsight one could state that failure of these two systems seemed unavoidable, this is not the case. Designing successful computer-based patient information systems, we claim, is a very complex task. In order for computer-based systems to be successful, designers have to make use of the specific advantages and limitations of IT. In addition, computer-based patient information systems also have to live up to users'—often high—expectations, partly caused by the rhetoric that is used to persuade patients to use systems.

1.1 Background

The first evaluation project was a project of the Rotterdam Eye Hospital [15]. The Rotterdam Eye Hospital is the only hospital in The Netherlands that is specifically oriented towards eye afflictions. One of the most common eye afflictions in children is amblyopia, in which normal vision in one eye fails to develop because of a difference in vision between the two eyes in early life. The affected eye is also called 'a lazy eye'. The hospital started a project in 1999 to improve the quality of care in terms of effectiveness, efficiency and patient centeredness for children with amblyopia and their parents (see Section 2 for a more detailed description). Amblyopia can only develop in very young children and treatment has to start as early as possible, but in any case before the child has reached the

age of 6 years. Treatment usually involves patching of the unaffected eye. To improve the quality of care the hospital has developed an interactive, computer-based patient information system directed at both children and parents. At the moment that the Internet was becoming a frequently used medium in The Netherlands, the Eye Hospital wanted to find out if using this medium for giving patient information could help improve the quality of care. They chose children with amblyopia and their parents as a target group, because amblyopia is a frequently occurring eye problem and because children and their parents were considered to be a population that were more commonly using computers than for example elderly people. Together with the Dutch Digital Hospital, an organization that develops information technology based communication and information tools, they developed a patient information system. The system consisted of a CD-ROM and an Internet site. The basic material on the CD-ROM had been developed a few years earlier by the Eye Hospital and contained information on the hospital itself, on amblyopia, on the investigations done to establish the diagnosis of amblyopia, on the possible results of these tests, on causes, consequences, treatment methods and possible complications. The information was presented by an orthoptist, an ophthalmologist and the child health center physician in brief video fragments. Other fragments showed amblyopic children and parents speaking about their experiences. In addition, a CD-ROM featured a cartoon about Paul, a boy with amblyopia, who wears glasses and an eye patch.

The Internet site, specifically developed for this project by the Dutch Digital Hospital, contained four parts: a Chat box, a Question and Answer section, a Newsletter and Games. The Chat box afforded virtual contact between the parents and/or the children, supervised by an orthoptist or an ophthalmologist. During the pilot evaluation phase, the users were able to use this facility one night a week, during 1 h. In the Question and Answer section, the users could ask questions to one another. The Newsletter gave general information of the Rotterdam Eye Hospital. The Games consisted of coloring pictures, simple computer games and jokes for children. Both the Internet site and the CD-ROM contained images and voice recordings that made them accessible for parents as well as children. The second evaluation project was a project of the association of general practitioners in Rotterdam (Districts Huisartsen Vereniging Rotterdam). This association is a regional organization of and for general practitioners and is associated with the national association of general practitioners (Landelijke Huisartsen Vereniging). Four general practitioners of this regional association wanted to know whether it was possible to improve the quality of care in terms

of effectiveness, efficiency and patient centeredness (see Section 2 for a more detailed description) by installing a patient information system ('Digidoc 2000') in the waiting room of the practice. They received financial support from the regional organization of general practitioners. A small IT company, together with the GPs, developed the patient information system. The patient information system consisted of a small desk with personal computer with keyboard, mouse and printer. The system had two purposes. First, it was meant to show three kinds of medical information: generic information on common diseases, more in-depth information on specific conditions and a medical encyclopedia. The generic information on common diseases was identical to the paper brochures that are often present in the cupboard in the waiting room or the consultation room. Such information deals with common complaints, treatments and questions, such as lower back pain and its possible remedies. The more in-depth information on diseases is information normally given to the patient during consultation. It contains more medical terminology and specific advice for the individual patient. The medical encyclopedia was designed to be used in an environment where patients could consult the system and also have the opportunity to look up information (books, journals) the system refers to. This system allowed the user to search for both general and more detailed information on diseases, including pictures and short movies. In addition to all this, the system was also to replace the sign where announcements were made about for example holidays and special consulting hours. For this purpose, the designer made a bar that scrolled continuously on the bottom of the screen and showed these announcements [16].

1.2 Goals of the evaluations

In both projects, the initiators asked the Department of Health Policy and Management of the Erasmus University Rotterdam to evaluate the effects of the system on the quality of care. Both projects served as a pilot for us to investigate the possibility and feasibility of a larger study, and for the initiators to decide how to 'scale up' the projects. The project for children with amblyopia and their parents was to be followed by a randomized controlled trial to prove the increased effectiveness, efficiency and patient centeredness that this system would bring. The project for patients in the waiting room of the general practitioner was to be followed by implementation in more practices of general practitioners in Rotterdam. We were interested in conducting these evaluation studies because exactly such explorative studies (small scaled and with a primary focus on qualitative methods) can teach us about the experiences

of patients with novel ways of employing information technology and the reasons for (not) using it.

2 Methods

The impact of information systems on the quality of care delivery is often measured primarily by looking at the effectiveness of the care provided. Here, quantitative evaluation approaches are the implicit golden standard [17]. For example, information systems designed to improve decision-making by doctors, so-called Clinical Decision Support Systems (CDSSs), are often evaluated by the amount of times that the system can identify the right diagnosis, or the rate in which it outperforms the decision-making of human experts [18]. Likewise, the effectiveness of for example automatic alerting systems may be measured by the reduction in time between the emergence of a critical laboratory result and the ordering of an appropriate treatment, or the difference in the number of adverse events with or without an information system [19].

In these pilot projects, however, both the initiators and us defined the quality of care delivery in a broader sense: including the dimensions of *effectiveness* (e.g. improved clinical outcomes), *efficiency* (e.g. fewer consultations) and *patient centeredness* [20]. In addition, we were keenly interested in the *nature* of the impact and in the question *why* the patient information systems had a particular impact. Therefore, qualitative and quantitative methods were combined (see next paragraph).

Research methods in the first project consisted of observation of orthoptists' and ophthalmologists' consultations, in-depth exploration of the functionality of the patient information system, virtual observation of chat sessions, in-depth interviews with users and observation of the system in actual setting of use (i.e. the patient's home). The researcher first observed several orthoptists' and ophthalmologists' consultations, to become familiar with the treatment. This was important in order to understand how parents and caregivers deal with amblyopia. What kinds of questions are asked (i.e. what kind of knowledge and worries do parents have) and what kind of answers are given. Exploration of the system made clear how it was set-up and how it worked. Virtual observation of chat sessions showed how long and how many people joined the sessions, which kind of questions they asked, to whom they preferred to talk and what kind of discussions ensued. Subsequently, one of the parents of each family was interviewed. The children appeared to be too shy or young to answer questions, so the parents informed the researcher about the experiences of their child with the project. Because of the

explorative nature of the study, no inclusion criteria regarding selection of the parents - apart from access to the system - were necessary. Finally, in two cases, the researcher observed the system in the actual setting of use (i.e. in the patients home). Rather than for example using questionnaires, we conducted observations and in-depth interviews. Successful use of information technology, after all, is dependent on many parameters. Just what kind of benefits the users will perceive is often difficult to predict [21]. The—subtle—interplay between characteristics of the institution (e.g. attitude and assumptions towards patient education), the user (e.g. interest in and need for electronic education) and the technology (e.g. interface, speed) are crucial to the fate of the system [14]. Hence, it was necessary to use methods that enable to analyze users' real-time experiences (observations) and make it possible to clarify respondents' answers (interviews). The interviews were done by means of a topic list that covered all the relevant aspects of the research questions: reason for participation, need, use and satisfaction regarding functionalities of the system, experienced differences with traditional patient education, general computer experience, experienced outcomes.

The qualitative research methods for the second project consisted of observations of the system in use in the waiting room, interviews with patients and interviews with the general practitioners and the secretaries. No patients were excluded, except for those that were too old or ill to ask questions. The researcher spent an average of 2.5 days of observations in the waiting rooms of each practice. The observations helped to understand the setting of the waiting room as an environment for patients to use the system, and were focused on the way the system was used, the interactions between patients that were waiting and patients that were using the system, and the location of the system in the waiting room. The interviews with the GPs and their secretaries served to find out whether they experienced any differences in patient behavior related to the use of the system, such as more or less questions during consultation (in case of the GP) or before or after consultation (in case of the secretary), more satisfaction because of the possibility to look up information, better informed patients and so forth.

Also quantitative measurements were taken. The system registered in a log file how many people used the system, which of the three sources of information they consulted, for how long, and whether they made a print of the information. In addition, every time one of the information sources was consulted, users were asked whether they had found what they were looking for and if they were satisfied with the information. They were also asked about their age, education and computer experience.

2.1 Measurements

Effectiveness was measured by the (1) actual use of the system, (2) reduction of anxiety and (3) increasing compliance. In the first project, actual use was measured by analyzing the use of the Chat box and the Question and Answer section, and by interviewing the parents. Actual use in the second project was measured by the observations and analysis of the log files. In the first project, reduction of anxiety was measured by analyzing the questions parents posed on the Chat box and the Question and Answer section, and by means of interviewing the parents. In the second project, reduction of anxiety was measured by means of interviewing the patients and GPs. Increasing compliance was measured by interviewing the parents (first project) and the GPs (second project). *Efficiency* was measured by changes in the amount (first project) and length (second project) of consultations (as subjectively estimated by physicians). *Patient centeredness* was measured by investigating three features of both patient information systems: the interactive nature of the systems, the accessibility of the systems and the user friendliness of the systems. The interactive nature of the system was measured by asking the users what their opinion was of the nature of the system (computer, CD-ROM), compared to traditional means like books/leaflets or a video. The accessibility and patient friendliness was measured by studying real-time use and interviewing the parents (first project) and observations and interviewing patients (second project).

3 Results

In the *first project*, 14 families participated, with a total of 15 children with amblyopia. The average age of the parents was 35 years; the mean age of the children was 4.9 years. Two of the 15 children were 'new' patients (being in treatment less than 3 months) at the moment of joining the project, the others had been in treatment for a longer period of time (with an average of 24 months). Most of the parents were highly educated: in 9 out of 14 families at least one of the parents had a college degree. The *second project* consisted of the participation of four GPs and their secretaries, solo practices as well as group practices. The practices were located in different sites of Rotterdam, including deprived areas, as to have different types of practices and patients. Ninety-six patients were interviewed after consulting their general practitioner. The respondents mainly consisted of women, 65 out of 96, and most of them were younger than 44 years old. Fifty out of 96 stated that they had computer experience.

The results of the two pilot projects turned out to be similar: contrary to the expectations and hopes of the physicians and the designers, the systems were hardly used. For example, most of the parents of the children with amblyopia checked the CD-ROM only once or twice. In addition, half of the parents used the Chat box, but not for chatting with other parents, but for consulting the orthoptist or ophthalmologist. The newsletter, the games and the video fragments were not used at all. The patient information system in the waiting room was also hardly used. Almost half of the patients did not even know what the function of the system was, despite the fact that it was announced in different ways (on the door of the waiting room, at the side of the desk of the system and on the screen itself). As a consequence, the systems had no effects on the before mentioned indicators of quality of care.

A core reason for these disappointing conclusions, we argue, was that there appeared to be a large gap between what users actually needed and wanted, and what the designers had *assumed* they needed and wanted. Since the designers had built their assumptions—which were often implicit—into the system, this gap resulted in the systems being rejected by the users. Based upon the literature and our analysis of these two evaluation studies, we have singled out three types of assumptions regarding the use of computer-based patient information that were problematic. The first concerns the *emotional and cognitive content* of the information that is given. The second concerns the *moment in the illness course* that the information is offered. The third concerns the *setting* in which the information is offered. As such assumptions are necessarily built into the design of every patient information system, we hope that our discussion can help prevent similar mismatches in the future.

3.1 The emotional and cognitive content of the information

The *emotional and cognitive content of the information* that was offered reflected the designers' expectations of users' information needs. The designers of the project for amblyopia assumed that amblyopia was seem by the children and the parents as a serious condition. Children were expected to be teased by other children because their eye had to be patched, and parents were expected to get awkward reactions from the environment. As a consequence, the type of information that was offered to children—a cartoon of a boy with the message that he should not feel excluded because of amblyopia—was focused on trying to make the child feel better.

These assumptions, however, appeared to be quite problematic. The children in the pilot project did not have many negative emotions about

their patched eye and most children were never teased. The age group of these children was younger than the designers seemed to have aimed at: the problem of 'teasing' usually occurs at a later age than between 2 and 4 years. In addition, many parents we interviewed stated that 'patched eyes' are so prevalent that most children do not feel excluded at all.

The content of the information that was offered to the parents also appeared to be not in accordance to their need. On the CD-ROM they could consult information on the causes of amblyopia, the treatment and possible problems. They could also make contact with other parents and eye specialists for asking questions. This however, was not information they were waiting for. First, the parents indicated that a child with amblyopia is not thought to be afflicted with a serious condition that impacts heavily on daily life. Also, they reported that they experienced no difficulties with the treatment of amblyopia. Children usually get used to wearing the patch rather fast, and then they become rapidly indifferent to it. And because of the above, they felt no need to chat about amblyopia. Some of them even felt this chat option to be rather overdone. In addition, the children could do almost nothing with the system, because it seemed to be designed for an age group older than the average age of the children who constituted the target group. The parents estimated that 7 years would be the minimal age for a child to do something with the system that would affect the child in any of the ways hoped for by the designers of the system. The system therefore failed on an emotional level (the parents nor the children experienced negative emotions regarding amblyopia or its treatment), as well as on a cognitive level (the parents knew enough about amblyopia and had no difficulties with the treatment whereas the CD-ROM was too complicated for children).

The problem regarding the second project consisted of the assumption that the general practitioners as an information source could more or less be 'replaced' by a computer-based patient information system. What the designers and the general practitioner did not realize, however, is that patients visit their GPs not so much to be *informed* but to be heard and reassured, to talk and to get explanations [22]. It is the *interaction* that counts. The information patients receive during their consultation emerges from this interaction, and is therefore geared towards this specific patient's worries, questions, intellectual capacities and so forth. This function, of course, cannot be easily fulfilled by a computer-based patient information system, as was also clearly stated by several patients in our research.

> 'I do not want to consult a computer; I want to talk to my general practitioner. How can I have a normal conversation with a computer?' (Parent A)

Apart from the emotional component of information, replacing the general practitioner by the computer-based patient information system assumes that patients know what kind of information they need to receive and that they are able to find and interpret this information by themselves. In practice, however, patients go to their general practitioner because they often do not know what their complaint is about and want an expert to look at it. Finding and interpreting the right medical information by themselves, as a consequence, often is problematic [23–25].

In addition, the designers and the general practitioners assumed that patients would appreciate computer-based information more than information on paper. They thought that since computer-based media, for example the Internet, are becoming more and more popular, people are more attracted to consult computer-based information than information on paper. This assumption, however, is dubious. Especially in the setting of the waiting room (see further), brochures and leaflets can be picked up easily, while computer-based information first has to be searched for and also has to be printed. Though computer-based presentation of information offers more possibilities (for example, regarding the amount of information that can be stored and retrieved and the use of search engines), it also makes it more complicated to find the right information. And the more complicated the presentation of information, the smaller the chance that the information is effective [26]. The people that consulted the encyclopedia, for example, experienced that the interface of the system was disorderly and that they had trouble finding the right keywords. When they found the right keywords, they were referred to information sources (like books) that were not available in the practice. So, this system also failed on an emotional level (information is not only about words on paper or on a screen, but also about interaction) as well as on a cognitive level (patients often do *not* know how to find and interpret the right medical information and do not by definition prefer computer-based information).

3.2 The moment in the illness course

The second problematic assumption in both projects was concerned with the *moment in the illness course* in which the (possibility to obtain) information is offered. By moment in the illness course we refer to the moment that the information is needed and suitable to comprehend. As

is known from the literature, different patients have different information needs, and the needs of particular patients vary over time. Not only do systems have to give explanations adapted to patients' gender and age, and possibly to their educational level and ethnic backgrounds as well, it also needs to be tailored to patients emotional concerns, including where individuals are in their understanding of and coping with their condition [27].

Because of the type of questions and worries patients have, face-to-face contact is therefore often the most suitable in the early phases of a patient's illness course. Patients' needs regarding information after this initial contact depend on the complexity of the disease in terms of treatment, unpredictability of the illness course, necessary lifestyle changes, impact on daily life, etc. In the case of amblyopia, it appeared that parents were confronted with the system relatively late after they learned about the diagnosis. Though the parents were a bit frightened in the beginning and uncertain about the consequences of the diagnosis, they stated that this did not last long. By following the instructions from the orthoptist, they soon realized that amblyopia is not a severe illness, that it is a temporarily affliction and that it can be treated well. So, when they were confronted with the system, between 4 and 89 months after treatment, their information needs were already fulfilled. Even patients that were in treatment fewer than 3 months stated that they did not get much support or information from the system. For them, information from the orthoptist, friends and relatives was sufficient.

> 'Actually, it was clear from the beginning on. At the first appointment you get a brochure and that contains a lot of information, so in fact I had enough from the start.' (Parent B)

So, first, the information given to patients on CD-ROM was too late: patients knew enough about amblyopia and had no need for further information. Second, though some parents were worried in the beginning, amblyopia is a relatively simple disease with a simple treatment with hardly any complications. Because of this, patients had no need for additional information provided by the computer-based patient information system.

For patients using the patient information system in the waiting room of their general practitioner, the timing of offering patients the possibility to look up information was also a problem. Patients normally look up information in the period *before* they consult their general practitioner [28]. Patients themselves also explicitly stated this. They stated that *if*

they wanted to have looked up information, they would have done that at home talking to their partner or using a medical encyclopedia or the Internet. The presence of the system even irritated a few patients, because they were afraid that their general practitioner would perhaps push them to use the system at the cost of consultation time.

3.3 The 'setting'

The third and final problematic assumption is concerned with the 'setting'. This term refers to the technical, physical and organizational context of use, such as limitations regarding access to the system or the characteristics of the place where the information system is used [29]. The setting does not determine what happens, but it constrains or shapes peoples' experiences and actions regarding the use of the system. When the parents of the children with amblyopia had specific questions, they had the possibility to chat to other parents or an orthoptist/ophthalmologist via the chat function on the CD-ROM. However, the only time slot where it was possible to chat for the parents, was between 19:00–20:00 h on Thursdays. This is problematic because this time is exactly 'rush hour' in their households.

> 'The timing of the chat sessions was horrible! It is exactly at the time that I feed them, wash them and bring them to bed.' (Parent C)

Also, parents felt it was very impractical that they could only ask the questions to their caregivers once a week. They would rather have been able to put the questions on a discussion list that can be consulted whenever they wanted.

The patients in the waiting room of the general practitioner appeared to be even more constrained by the environment where the system was set-up. The idea was that the waiting room is a suitable environment for patients to consult medical information because it was expected that most patients like to look up medical information. Also, the idea was that patients, while waiting for their turn, would like to do something to kill time. However, the designers and the general practitioner had not paid attention to the fact that, as appeared from our study, the waiting room is a very specific type of semi-public place. First, the system was meant to offer the possibility to consult medical information *and* to give general information about holidays, special consultation hours, etc. to *all* patients. Therefore, the system was placed in such a way that people could see the 17 inch. screen from practically all positions in the waiting

room. This, of course, hampered the use of it. Especially for sensitive topics, this is a problem [30].

> 'I do not feel like looking up information, while the other people in the waiting room can see what I'm looking for.' (Parent D)

Second and more important, the waiting room, sociologically speaking, is a very specific type of place where most patients are actively 'not doing anything'. They just sit and look around them. Some patients take a magazine or a newspaper, but only seem to glance through it. A waiting room is different from for example a living room, a library or the consultation room of the general practitioner—it is more like an elevator or another confined space in which (relative) strangers find each other, waiting together in close proximity. In all of these specific types of places certain behavior is stimulated, expected, constrained, etc. and therefore seen as 'normal' or not. In the (often small) waiting rooms of general practitioners, calling with a cell phone, for example, is not seen as 'normal'. From the observations, it appeared that even walking through the waiting room and taking brochures or leaflets from a cupboard attracts attention from the other patients: all the eyes turn towards the individual 'breaching' the atmosphere of 'active inactivity'. So, though the waiting room, in practical sense, might seem a suitable room because patients have time to look up and read patient information material, in practice this is not the case. And it is exactly because of this atmosphere that patients, when consulting the computer-based patient information system, break with 'the spirit or ethos of the situation'. 'The rule of behavior that seems to be common to all situations (…) is the rule obliging participants to 'fit in" [31]. Consulting the system in an environment where patients do not really *do* things like reading or talking, means finding oneself in a very visible and therefore often awkward position.

4 Discussion and conclusion

In this chapter, we discussed two computer-based patient information systems. Both systems were meant to replace already existing, traditional means of patient education such as brochures, leaflets and to a certain extent face-to-face communication. Both systems, however, were hardly used; when they were used, the patients were predominantly dissatisfied. In our analysis, we showed that the information systems failed because the emotional and cognitive content of the information was not geared

towards the actual needs of the patients. Also, the moment in the illness course and the setting in which information was offered, were problematic.

It could be argued, perhaps, that one could have predicted beforehand that both projects were bound to fail. Yet, this would underestimate the complexity and the unpredictability of success and failure of IT applications. The assumptions of the designers were, on face value, not so strange: the participants in both projects broadly shared them. Because it is often difficult to predict what future users 'need' or 'want', the use of computer-based information systems for patient education requires a thorough analysis of the specific advantages and limitations of the different means of patient education. Interestingly, however, such comparative questions are not often posed: instead, most research is done to investigate how 'traditional' means of patient education can be replaced by computer-based patient information systems [4,23,32–34]. This starting point, however, is strange considering the differences in the strength and weakness of the different means for patient education. For example, face-to-face information is very personal and (potentially) specific, but also difficult to remember for patients and time consuming for the health professional [35,36]. Paper-based flyers, brochures and videos are impersonal, often unspecific but easy to make, distribute and share. Looking at all these differences, no single means can be the best in all the different situations. Face-to-face patient education will remain, at least for the time being, the most personal and interactive way of educating patients, where both verbal and non-verbal means of communication can be fully employed. On the other hand, it is also known that patients only remember a small part of the things that are said to them during consultation. Brochures make it possible to read text over and over again, in practically all situations one can think of, but are often very general and unspecific. The Internet makes it possible for patients to search for information and contact each other all over the world, from their own homes on the moments they choose to do that. The Internet also makes it possible for patients to ask doctors a question, without having to go to the doctor or hear the answer at the same time. But, the Internet is not accessible for all patients and does (as yet) not give the opportunity to verify on a real-time basis whether patients understand the information that is given. Consequently, the different means should be used in those circumstances where they gain their optimum impact. Looking at the project for parents and children with amblyopia, patient education was especially important immediately after the diagnosis was made. The parents appeared to be worried about consequences since they did not

know very well what to expect. Patient education, in this example, should therefore focus on informing and comforting patients by face-to-face education and brochures that explain what amblyopia is about. In this case, brochures are suitable because the information on amblyopia and its treatment is limited, straightforward, easy to understand and with relatively little impact on parents or children. Since the initial treatment at the hospital consisted, in fact, of face-to-face education and brochures, the information system was superfluous. Also, parents appreciated the opportunity to ask questions to specialists but since they were not worried that much, e-mail contact, for example, would have been a better option than a chat hour once a week [37].

The project of the patient information system in the waiting room of the general practitioner was problematic because of both the moment in time and the setting in which the information was offered. For the kind of information that the system provided, it seems much more feasible to refer patients to the website of the general practitioner, where patients can look up information about the practice and about illnesses, for example with links to reliable web sites. Patients that want to look up medical information, then, are able to consult information on the moment they choose to do that, at their own pace and from the environment of their choosing. (Some Dutch general practitioners are starting to do just that.)

Building successful computer-based patient information systems is not only a complex task because of the reasons mentioned above. Information technology, the Internet in general and health care IT applications specifically, are often mentioned as *the* solution for the limitations of 'traditional' means for patient education that we mentioned in the introduction. Users, for example, are promised full access to their own medical files, online consultation or medical databases that make it possible to look up all kinds of information by using the right keywords [38].

As a consequence, the expectations of users are often very high. Designers have to live up to these expectations, also because users often have to invest a lot of effort in learning to work with IT applications. This means that computer-based patient information systems have to have serious additional advantages compared to traditional means of patient education. In both cases, however, the designers trusted too much on the fact that they offered patients something new, without investigating the actual surplus value for patients. However, projects like these that are too much technology driven instead of based on thorough analysis of advantages and limitations of IT compared to traditional means, are bound to fail [21].

4.1 Conclusion

To conclude, making use of the full potential of advantages of IT in patient education still seems a long way ahead. However, there are patient information systems that have proven to be successful. Looking at the literature, especially systems for patients with chronic diseases like diabetes, diseases with a sensitive nature that benefit from privacy in communication or diseases in which simplified language and self-paced instruction is of an advantage, appear to be successful [3,39]. Using IT in this way requires dedicated systems, patients and health-care professionals. But when these conditions are fulfilled, computer-based patient information systems are able to improve health status and a valuable supplement to (rather than a substitute for) traditional means of patient education [39–41].

4.2 Practice implications

- Designers of patient information systems should be aware of their assumptions regarding patients' needs regarding the use of information technology for patient education and should therefore investigate thoroughly whether these assumptions indeed match real patients' needs, desires and possibilities. To investigate whether it is worthwhile to further develop a design, one can often use - relatively simple - research methods: interviewing a random sample of patients, observations of the setting in which the system is planned to be used, and studying comparable evaluation studies.
- Strategies to *replace* traditional means of patient education by computer-based patient information systems often do not do justice to the complexity of patient education and the individual weaknesses and strengths of the different means for patient education. Instead of a strategy of replacement, computer-based patient information systems should be seen as a potentially valuable *supplement* rather than a substitute, whose specific strengths and weaknesses will have to be matched anew to patients' needs, desires and/or possibilities.

References

1 Silva JS, Zawilski AJ. The health care professional's workstation: its functional components and user impact. In: Ball MJ, Collen MF, editors. Aspects of the computer-based patient record. New York: Springer; 1992. p. 103–23.
2 Mead V, Rhyne R, Wiese W, Lambert L, Skipper B. Impact of environmental patient education on preventive medicine practices. J Fam Pract 1995;40:363–9.
3 Lehmann ED. Interactive educational simulators in diabetes care. Med Inform 1997;22:47–76.
4 Skinner CS, Siegfried JC, Kegler MC, Stretcher VJ. The potential of computers in patient education. Patient Educ Couns 1993;22:27–34.
5. Bental DS, Cawsey A, Jones R. Patient information systems that tailor to the individual. Patient Educ Couns 1999;36:171–80.
6. Dijkstra A, De Vries H. The development of computer-generated tailored interventions. Patient Educ Couns 1999;36:193–203.
7. Brennan P, Strombom I. Improving health care by understanding patient preferences: the role of computer technology. JAMIA 1998; 5:257–62.
8. Tang P, Newcomb C. Informing patients: a guide for providing patient health information. JAMIA 1998;5:563–70.
9. Brug J, Campbell M, Van Assema P. The application and impact of computer-generated personalized nutrition education: a review of the literature. Patient Educ Couns 1999;36:145–56.
10. Lewis D. Computer-based approaches to patient education: a review of the literature. JAMIA 1999;6:272–82.
11. Lennox AS, Osman LM, Reiter E, Robertson R, Friend J, McCann I, et al. The cost effectiveness of computer tailored and non-tailored smoking cessation letters in general practice: a controlled trial. BMJ 2001;331:1396–400.
12. Forsythe DE. New bottles, old wine: hidden cultural assumptions in a computerized explanation system for migraine sufferers. Med Anthropol Q 1996;10:551–74.
13. Forsythe DE. Using ethnography in the design of an explanation System. Expert Syst Appl 1995;8:403–17.
14. Street J, Richard L. Mediated consumer–provider communication in cancer care: the empowering potential of new technologies. Patient Educ Couns 2003;50(1):99–104.
15. Van't Riet A, Berg M, Hiddema F, Sol K. Meeting patients' needs with patient information systems: potential benefits of qualitative research methods. Intern J Med Inform 2001;64:1–14.

16. Stoop AP, Berg M. Elektronische informatie in de wachtkamer van de huisarts (Electronic information in the waiting room of the general practitioner). Huisarts Wet 2001;44:101–2.
17. Kaplan B. Evaluating informatics applications—some alternative approaches: theory, social interactionism, and call for methodological pluralism. Intern J Med Inform 2001;64:39–56.
18. Friedman CP, Elstein AE, Wolf FM, Murphy GC, Franz TM, Heckerling PS, et al. Enhancement of clinicians diagnostic reasoning by computer-based consultation: a multisite study of 2 systems. JAMA 1992;82(19).
19. Kuperman GJ, Teich JM, Tanasijevic MJ, Luf NM, Rittenberg E, Jha A, et al. Improving response to critical laboratory results with automation: results of a randomized controlled trial. JAMIA 1999;6:512–22.
20. Stoop AP, Berg M. Integrating quantitative and qualitative methods in patient care information system evaluation: guidance for the organizational decision maker. Methods Inf Med 2003 [in press].
21. Berg M. Patient care information systems and health care work: a sociotechnical approach. Intern J Med Inform 1999;55:87–101.
22. Grol R, Wensing M, Mainz J, Ferreira P. Patients' priorities with respect to general practice care: an international comparison. J Fam Pract 1999;16:4–11.
23. Richards B, Colman AW, Hollingsworth RA. The current and future role of the Internet in patient education. Int J Med Inform 1998;50:279–85.
24. Rose S, Bruse J, Maffulli N. Accessing the Internet for patient information about orthopedics. JAMA 1998;280:1309.
25. Van der Kar A, Knotnerus A, Meertens R, Dubois V, Kok G. Why do patients consult the general practitioner? Determinants of their decision. Br J Gen Pract 1992;361:313–6.
26. Wapenaar H, Roling NG, Van den Ban AW. Basisboek voorlichtingskunde (Handbook on patient information). Amsterdam: Boom Meppel; 1989.
27. Forsythe DE. Using ethnography to build a working system: rethinking basic design assumptions. Proc Annu Symp Comput Appl Med Care 1992:505–9.
28. Schepers RMJ, Nievaard AC. Ziekte en zorg: inleiding in de medische sociologie (Health and lllness: an introduction to medical sociology). Houten: Educatieve Partners Nederland BV; 1997.
29. Silverman D. Communication and medical practice. Social relations in the clinic. London: Sage; 1987.

30. Ward K, Hawthorne K. Do patients read health promotion posters in the waiting room? A study in one general practice. Br Med J 1994;44:583–5.
31. Goffman E. Behavior in public places. London: Collier-Macmillan Ltd.; 1963.
32. Rostom A, O'Connor A, Tugwell P, Wells G. A randomized trial of a computerized versus an audio-booklet decision aid for women considering post-menopausal hormone replacement therapy. Patient Educ Couns 2002;46:67–74.
33. Dragone M, Bush P, Jones J, Bearison D, Kamani S. Development and evaluation of an interactive CD-ROM for children with leukemia and their families. Patient Educ Couns 2002;46:297–307.
34. Jones JM, Nyhof-Young J, Friedman A, Catton P. More than just a pamphlet: developments of an innovative computer-based education program for cancer patients. Patient Educ Couns 2001;44:271–81.
35. Calkins DR, Davis RB, Reiley P, Phillips RS, Pineo KL, Delbanco TL, et al. Patient–physician communication at hospital discharge and patients' understanding of the post-discharge treatment plan. Arch Intern Med 1997;157:1026–39.
36. Tang PC, Newcomb C. Informing patients: a guide for providing patient health information. JAMIA 1997;5:563–70.
37. Ferguson T. Online patient-helpers and physicians working together: a new partnership for high quality health care. Br Med J 2000;321:1129–32.
38. Kim M, Johnson KB. Personal health records: evaluation of functionality and utility. JAMIA 2002;9:171–80.
39. Krishna S, Balas EA, Spencer DC, Griffin JZ, Boren SA. Clinical trials of interactive computerized patient education: implications for family practice. J Fam Pract 1997;45(1):25–33.
40. Strömberg A, Ahlén H, Fridlund B, Dahlström U. Interactive education on CD-ROM—a new tool in the education of heart failure patients. Patient Educ Couns 2002;46:75–81.
41. Plougmann S, Hejlesen OK, Cavan DA. Dias-Net—a diabetes advisory system for communication and education via the Internet. Int J Med Inform 2001;64:319–30.

Chapter 4

Stoop AP, Vrangbæk K, Berg M. Theory and practice of waiting time data as a performance indicator in health care. A case study from The Netherlands. *Health Policy* 2005;73:41-51.

Abstract

In this study we investigate the use of waiting time data as a performance indicator in health care in The Netherlands. We explain why the current publication of waiting time data fails to achieve one of the main goals: to have consumers and general practitioners act upon this information. The reason, we claim, is that even seemingly clear-cut, easily measurable and objective numbers such as waiting times need interpretation to become meaningful. Discussing four themes – the patient behind the number, the treatment behind the number, the strategy behind the number, and the specificity of the number – we will discuss just how deeply this need for interpretation affects the usability of 'waiting times' for purposes such as informing consumers. We will argue that this problem is due to not making a clear distinction between performance indicators for internal use and for external use. We conclude that the usefulness of the publication of waiting time data for consumers strongly increases when waiting times are guaranteed and related to treatment options like booking possibilities and other performance indicators such as patient satisfaction.

Keywords: Health care; The Netherlands; Performance indicator

Theory and practice of waiting time data as a performance indicator in health care. A case study from The Netherlands

Arjen Stoop, Karsten Vrangbæk, Marc Berg

1 Introduction

The public presentation of health care performance information remains much debated. There is increasing consensus about the fact that consumers[4] should be well informed about the quality and characteristics of care of different healthcare providers to be able to make conscious decisions about their care [1,2]. Healthcare performance data have been made public for more than a decade in the United States and England and the production and dissemination of performance data is now a multimillion-dollar industry. Yet the impacts are doubtful. It has proven difficult to judge whether the benefits of public disclosure of performance data outweigh the disadvantages such as costs and administrative burden for health personnel [3].

One of the most common assumptions about performance indicators (PI) is that the public presentation of performance data is a prerequisite for the development of a properly functioning health care *market*. Information about performance and costs should encourage consumers to choose those providers that offer high-quality service at a low cost [3]. Evidence suggests however that performance data only have limited impact on consumer decision-making. In a review article Marshall et al. found only one article indicating that consumers take performance data into account when taking healthcare decisions [3]. The reasons for consumers' lack of interest in and use of performance data include difficulties in understanding the information, lack of interest in the nature of the information available, lack of trust in the data, problems with timely access to the information and lack of choice [3]. In addition, there is a growing awareness that there is little knowledge about what

[4] Patients and consumers are used as synonyms in this paper.

information people actually want or need to make their choices [4,5]. Currently, 'report cards', 'provider profiles' or 'consumer reports' are being produced that are explicitly designed to help consumers make better informed healthcare decisions by disseminating comparative information on price, access and quality of care. Yet even in such cases, consumers hardly act upon the performance information [6,7].

In addition, information about performance is also meant to inform other actors in health care. General practitioners, since they refer patients to the hospital, are important in helping the patient choose the right doctor or hospital. Insurance companies and public 'purchasing' organizations are able to contract those hospitals that offer the best quality of care. However, evidence that these actors do act upon published performance information is similarly lacking [8].

Why do consumers not act more extensively upon performance data when choosing a doctor or a hospital? And why do payers or referring professionals not use performance data? To answer these questions we will analyze in detail one specific performance indicator: waiting time. As is the case with many other performance indicators, waiting times are published with the intention to influence patients, general practitioners and health insurance companies to select 'high quality care'. As is also the case with many other performance indicators, however, these actors did not appear to act upon published waiting time information [9–11].

In this study we refer to waiting time as the time patients have to wait on a waiting list to be eligible for 'standard' operations like eye lens operations, knee replacements, hip replacements, hernia operations, etc. We do not refer to the time patients have to wait for their turn in the clinic itself, e.g. for an X-ray or MRI scan at the radiology department. We also do not refer to waiting times for special operations like organ transplantations. Though long waiting times are often associated with long waiting lists, this does not need to be the case. However, where the authors or the respondents mention waiting lists, they are referring to waiting lists with long waiting times.

Though the reasons for the existence of waiting times of the above mentioned 'standard' operations might differ from hospital to hospital, in general waiting times are seen as being caused by a sub-optimal circulation of patients. Waiting times are therefore seen as a (process) performance indicator of the organization. Analyzing the way waiting times are interpreted and used offers therefore a good opportunity for showing pitfalls, strengths and weaknesses of the publication of PI.

2 Methods

In order to find out *why* patients and general practitioners hardly act upon the publication of waiting time data and which potential problems these actors have in interpreting waiting times, we conducted a process evaluation on the way waiting time data were generated, interpreted and used in 26 hospitals in five regions in The Netherlands. These regions had been experimenting several years (between 1993 and 2001) with generating and publishing waiting times, prior to the national obligation for hospitals to publish their waiting times. Since we realized that the topic of waiting time data is strongly connected to all kinds of (professional, political and financial) interests, we decided to focus on interviews with those actors that were explicitly involved in these regional projects: hospital management, GPs, health insurance companies, consulting physicians and representatives from patient organizations. We conducted 35 semi-structured interviews. The topic lists were based on the minutes of the working groups, which were obtained by the chairs of the working groups. Analyzing the minutes of the working groups enabled us, first, to grasp the progress of the working groups and to identify the actors' formal role in this. Second, by analyzing this, the minutes offered the opportunity to get an idea of the actors' interests and position regarding waiting time data. Since the regions differed in the start date of the project, the formation of the working group, the initiator, the media for publishing waiting time data, the way they made progress and the problems that were encountered, our topic list differed for the different regions. We focused on the actors' expectations, aims, interests and encountered problems regarding the publication of waiting times. By doing this, we tried to find out what waiting times *mean* to them [8]. All the interviews were transcribed and analyzed by using ATLAS.ti.

3 Interpreting waiting time data

If waiting times are meant to inform patients, general practitioners and third party payers, the question is subsequently: what is these actors' opinion on waiting times and (how) do they use waiting times? For interpreting waiting times, we will show, four different themes need to be analyzed. By showing how waiting times are connected to these themes, we illustrate the difficulty of interpreting and using waiting times for those actors – such as patients, general practitioners and third party payers – that are not aware of the context in which waiting times are generated. After this analysis, we will zoom in again on the problems

and potentials of the publication of waiting data time specifically for patients and general practitioners.

3.1 The patient behind the number: the significance of waiting times

In the media, but also in scientific journals, waiting times are often discussed in terms of negative consequences, like the risk of deterioration of the patient's condition. Yet is waiting for an operation, for example, problematic for all patients? And if not, for who is it a problem? The mere *existence* of a waiting time, several respondents claim, cannot be interpreted as a problem without knowing more specific details of the individual patient 'behind the number'. Though the waiting time for cataract and hips and knee replacements in The Netherlands is generally seen as problematic because of its length, a manager of a hospital explains:

> 'My opinion is that the waiting times for hip-and knee replacement and cataract are not by definition problematic. I mean, to wait half a year for a new lens seems like a very long time, but this is very subjective. Some eye doctors tell patients that have cataract that they need a new lens as soon as possible. Others might say to the same patient, let us wait another year. This is the same with a new hip or knee.'

Because doctors use different criteria and therefore diagnose differently, and since this is seldom explicated, it is hard to say – in general – how problematic it is for patients to wait. Another example provides a similar insight. The waiting times are generally long for consulting a cardiologist in The Netherlands. According to a cardiologist, the question is whether these long waiting times should be interpreted as a problem and whether all patients that are waiting should be helped within short notice.

> 'There are 1785 people on the waiting lists for cardiology all over The Netherlands. But such a number is hardly informative. If you want to be able to interpret that number in terms of seriousness, you have to have more data. For example, does the patient have a lot of complaints, how much medication does the patient get, how old is the patient? If you remove the people older than 80 years old from the list, how many are there left? This leads to a different number, and to a different interpretation.'

The problem about the significance of these waiting times is that they do not differentiate between patients. They do not distinguish between the

severities of individual cases. Though hospitals often informally prioritize patients to a certain extent, different kinds of patients (patients that are in pain or not, patients that are old or not, patients that want to be treated fast or not, patients that are distressed or not, etc.) are put on the same waiting list, without a specific order. Whether it is a problem for these patients to wait, depends on their individual situation and cannot be generalized, nor deduced from the number of waiting patients as such [12].

3.2 The treatment behind the number: the quality of care related to waiting times

Waiting times limit or restrict access to health care and are therefore an important concern for those interested in quality improvement. Waiting times are common in countries where resources are limited [12]. Very little is known, however, about the relationship between waiting times and the quality of care that is given by a health care organization or individual health care professionals. For example, long waiting times might reflect the popularity of a highly esteemed doctor, but could also reflect problems within the organization resulting in sub-optimal quality of care. Patients, just like customers looking at long lines in front of a restaurant, are sensitive to these kinds of 'cues', even though they may not know how to interpret them. A manager of a patient organization:

> 'The interpretation of the data is the problem. Long waiting times can be caused by an organization that is functioning badly and, as a consequence, might offer low quality of care, but also by the temporarily illness of a doctor. You have to be able to interpret that if you want to inform patients.'

Some respondents indicated that in their case the quality of care seems to be related to the waiting times. A cardiologist says:

> 'My perception is that our waiting times are strongly related to the contacts that one of our cardiologist has and the quality of care that we give. We have very good results and that is known far outside of this region. As a consequence, we get many patients from other places.'

However to 'outsiders', i.e. other actors than health care professionals, the relation between waiting times and quality of care is often unclear and hard to interpret. A manager of a regional organization for general practitioners about waiting times:

> "Some consulting physicians have a waiting time of three weeks, and others have a waiting time of 10 weeks. It does not give you any information. It does not mean that the one physician is doing his work badly, and the other is doing it well. That remains completely unclear. Waiting times are just numbers. What the background is? I do not know. Wards that have been closed, problems with personnel or consulting physicians that are of the opinion that patients just have to wait a bit longer. Once again, waiting times can have many causes."

Though patients, as well, are in general not aware of the possible relation between quality of care and differences in the length of waiting times, they tend to interpret the length of waiting times. As the manager of a department of a health insurance company states:

> 'Patients often ask the question why a certain consulting physician has a short waiting time and whether that is a bad sign. In that respect, patients ask more and more about the quality of care, whether it is a good doctor or not. I always say that there is no one-on-one relation between a long waiting time and a good doctor, and a short waiting time and a bad doctor. It is not like that.'

3.3 The strategy behind the number: waiting times and interests

Different actors in health care can have different and sometimes conflicting interests in maintaining or decreasing their waiting times [13]. For example, hospitals, consulting physicians and third party payers like health insurance companies are all interested in optimally utilizing their capacity and governing patient groups. Patients on the other hand, have other interests regarding their care (see further). Decreasing waiting times by tightening indications for access reduces costs and is in the interest of third party payers, for example, but not in that of the patient that wants to be referred (and is excluded for this specific indication). Establishing financial barriers for the patient, similarly, is also usually in the interest of health insurance companies or other payers, but usually not in the interests of the patient, physician or health organization. Patients might feel that they cannot afford care, which leads to a reduction of workload for professionals and organizations. The differences in interests between hospitals were also noticed clearly in the working groups. As one of the chairs from the working groups stated:

> 'We knew there was a risk of manipulation of waiting time data. You cannot control whether the numbers are true. If hospitals do not want certain patient populations, then they will not get them.'

One way of controlling the amount of patients on the waiting list is to limit the length of it. As a manager of one of the health insurance companies states:

> 'It will always be possible to manipulate the waiting times. For example, when do you put patients on the waiting list? There are consulting physicians that say that waiting times of six weeks are long enough. That means that new patients will be refused or referred to other consulting physicians or other hospitals. As a consequence, some consulting physicians have a waiting time of six weeks and others of twelve weeks.'

Another strategic aspect, directly related to money, is the length of waiting times. For many years, the Dutch Government had an action program against waiting times that allowed health institutions to get subsidies to counter long waiting times. This meant that having long waiting times was a way to get additional funding. Though the rules have changed, hospitals can still negotiate for more money if they can prove that they have long waiting times. This system, according to a manager of a health insurance company, should change.

> 'We still have a system that rewards those hospitals that have the longest waiting times. We have to get rid of this system at once.'

The above-mentioned differences in interests regarding waiting times reflect that waiting times are a political issue that involves all actors in health care: third party payers, health care professionals, hospitals, general practitioners and patients. As a consequence, it can be very difficult to disentangle the numbers that are published from the intentions, interests and practices that lay behind them.

3.4 The specificity of the number: the validity of waiting times

The fourth and last theme we want to discuss is the specificity and validity of the number. We will show how the process of generating waiting time data, the variations in work processes between hospitals, and the specificity of what is being waited for is defined, all limit the validity of the published waiting times.

It appeared that most Dutch hospitals had all kinds of problems generating waiting times electronically. The connection between the Hospital Information System (HIS) and the software application, for example, appeared problematic. The application itself also appeared to have problems. It showed different results in a row, even when the same analysis was done. As a consequence, practically all hospitals had to correct the numbers manually which required much time every month. And because updating the old waiting times also takes some time for Prismant (the organization that processes and publishes the waiting times on the internet), the data were never really up-to-date when they were finally published. Another problem was that some of the respondents had doubts regarding the numbers generated by the hospitals. As one of the policy makers of a patient organization states:

> 'Some hospitals have the same waiting times every month. This does not mean they are not reliable, but it is strange. You can also see on the website that some hospitals deliver new numbers every month and others do not.'

In addition, since hospitals have different ways of working, they also have different ways of generating their waiting times. Despite the fact that they all have to use the software application, there is still space for applying their own rules. One of the staff executives of a hospital elaborated:

> 'In the beginning we were obliged to use retrospectively generated data, but this has changed. All the hospitals are now allowed to work on the basis of expected waiting times. This means that there is space for making up your own rules. As a consequence, our waiting times cannot be compared with those of the other hospitals in the region, because I know that they are generated in a different way.'

Differences in the way hospitals work (because of differences in amount of specialties, personnel, etc.) are especially apparent between big (university) hospitals, and small (non teaching) hospitals. The differences between these two types of hospitals resulted in the fact that it appeared too difficult to generate comparable data, as one of the chairs of the working groups stated:

> 'Well, it appeared that because of the differences, the academic hospitals were not able to deliver the right numbers. This means that the

> numbers of the small hospitals are more valid than those of the academic hospitals.'

Not only between hospitals, but also within hospitals that are spread out over several locations, there are problems generating valid waiting times. One of the secretaries of a big hospital:

> 'We do not generate waiting times per location, but per consulting physician. In addition, consulting physicians sometimes have outpatient clinics on several locations, and those numbers are split up. This is confusing, because in the end all numbers are per specialty. So if you have to see a certain consulting physician, the numbers become unrecognizable and therefore unusable.'

The final issue regarding this theme has to do with what the patient is waiting for. As several respondents stated, waiting time data have a seeming precision that is not always in accordance with reality. How detailed should the differentiation in waiting times be [14]? As a manager of a hospital stated:

> 'Take for example surgery. There are different interventions and subspecialties in this discipline. What is the use of knowing that the shortest waiting time for a whole specialty is 10 weeks? In practice, the waiting time for an appendicitis is less than a week when it is urgent, and you have to wait six-months for a different, more complicated operation. The numbers do not mean a thing if you do not specify.'

But specifying waiting time data more precisely paradoxically increases problems of interpretation, a team coordinator of a hospital stated:

> 'There are six consulting physicians within internal medicine that all have their own specialization. Some physicians only have one-week waiting time, others have six weeks. So the question is, which waiting time should be published? When the referral is not directly related to one of these specialties, the patient nor the general practitioner knows in advance what the waiting time of that specific consulting physician is.'

For some specific operations, like hernias, hospitals are obliged to deliver waiting time data. But even two patients that both have to have a hernia operation can be very different with regard to the amount of time it

takes, the complexity of the operation, the required skills of the personnel, with different waiting time as a result. An interim advisor stated:

> 'We are obliged to deliver the waiting times for a hernia, but every single hernia is different. So publishing general waiting times for a hernia operation is impossible. I believe we still do not deliver those numbers.'

The gallbladder operation is another example of a type of operation that can be done in different ways with different waiting times as a result. However, these differences remain unclear in the registration. A project member:

> 'The gallbladder is in the registration system only known as a classic gallbladder. The gallbladder operation that is done with a laparoscope [an instrument that is used to operate in the abdomen via a little incision, AS] can not be found in this system.'

To summarize, analyzing these different themes leads to different kinds of interpretation problems, as can be seen in Table 2.

Table 2
The different themes connected to waiting times and their reason for the difficulty of interpretation.

Themes connected to waiting times	**Difficulties in interpreting waiting times for patients and general practitioners**
Theme 1: The patient behind the number: the significance of waiting times	Data lack sensitivity for the characteristics of the individual patient (e.g. age, medication, well being, preferences regarding (moment of) treatment).
Theme 2: The treatment behind the number: the quality of care related to waiting times	Lack of knowledge of the relation between the quality of care of a hospital, doctor or treatment and the corresponding waiting list.
Theme 3: The strategy behind the number: waiting times and different interests	It is unclear which intention/interest hospitals and consulting physicians have by publishing their waiting times.
Theme 4: The specificity of the number: the validity of waiting times	It is unclear how valid waiting times are due to the process of updating waiting times, differences in work practices between hospitals and problems of specifying treatments.

These four themes make it – especially – difficult for patients and general practitioners to interpret and act upon the data. For example, the first theme shows how difficult it is for patients and general practitioners to

judge the patient's situation and whether it is dangerous to wait this long for treatment or an operation? The data are just not transparent enough to make this kind of judgment. The second theme shows how unclear it is for patients what the relationship is between the length of the waiting time and the quality of care of that consulting physician or hospital, with the risk of being treated fast but sub-optimally. The third theme points at the strategy of waiting time generation and publication. Patients and general practitioners know that these data are being published with different intentions and serve multiple purposes, of which informing patients is not the only one. This obviously leads to reluctance in acting upon waiting times. The fourth theme discusses the process of generating waiting time data, the variations in work processes between hospitals, and what exactly patients are waiting for. Though one can choose for the hospital with the lowest waiting times on one specific group of operations (e.g. knee operations), it can very well be that for this patient's particular knee operation the waiting time is much longer.

Does the fact that patients and general practitioners do not use waiting times mean that publishing waiting times is useless to them and the other mentioned actors in health care? In the next section we will show that this is not the case. However, the usefulness of publishing waiting time data is dependent on the user and the goal of use. This match is crucial in order to make waiting times meaningful. Providing this match, however, requires an understanding of the difference between the use of performance indicators – like waiting time data – for internal organizational use (as an internal indicator) and for external public use (as an external indicator) [15].

3.5 Waiting time data as internal indicator

In the literature one can find different objectives of performance indicators (see the introduction). Some authors refer to internal use, i.e. for management purposes and quality improvement within the organization. Organizations that want to improve their care processes need indicators that make it possible to identify and analyze bottlenecks and to check whether changes lead to improvements. An important characteristic of internal indicators is that they refer to professionals' and managers' own context with, often, none or few changing variables in time (e.g. regarding patient population, treatment facilities, personnel, etc.). This means that relatively fast, small scale before-and after studies are sufficient and suitable to measure improvements. And because the professionals and managers involved are fully aware of the context in which the data are generated, they are par excellence able to interpret their own data. For

example, hospitals are aware of illness or shortage of personnel, financial problems that force them to focus on 'cheap' operations and postpone expensive operations, etc. And because hospitals are aware of this context, they are able to interpret the waiting times and use them for different internal, specific purposes. An executive secretary:

> 'From inside the organization we are very well able to see where the waiting times increase and what the causes are. For example, the last few weeks we analyzed the waiting times of varicose vein patients. By analyzing the way patients move through the hospital we are able to see which specialty does not optimally utilize the operation room capacity that is assigned. Capacity that is not used can, subsequently, be given to specialties that have more need for it. In this way we try to optimize our care processes internally.'

This use of waiting times, as internal indicator, appears to be very meaningful. Hospitals are able to analyze and improve their care processes by generating and acting upon their own waiting times. Importantly, this use of waiting time data is meant for the organization itself, not for 'outsiders' like patients, general practitioners or third party payers. And because of this internal use, no excessive precision or validation is needed since the numbers will not be used to make a comparison between health-care professionals or hospitals. This difference between internal and external use is crucial. Actors outside the hospital setting – especially patients and general practitioners – are interested in comparable data and therefore need to be sure that the waiting times are valid and precise. Actors within the hospital setting are aware of the context in which the waiting times are generated, which makes it possible to interpret them without validation or comparability. And since waiting times only have a specific meaning to those that are aware of the context in which waiting times are generated, they are neither relevant nor useful for general public like patients [16]. Stronger still, publishing waiting times that only make sense within the context in which they are generated, can be confusing and misleading to patients.

3.6 Waiting time data as external indicator

Instead of self-government and improvement, external indicators focus on enabling a comparison between health-care organizations or health-care professionals. External indicators are meant to inform the public in general and other actors (public authorities and health insurance companies) in particular. And to make a comparison, data have to be as precise

and valid as possible. Generating these indicators often requires time-consuming research, relatively large samples and aggregation and correction for all possible relevant differences between different health care settings. Since this is very complicated, those indicators that are able to fulfill these conditions are those that are measurable, and need not be those that are relevant [16]. A well known example of such a (clinical) indicator – meant to inform cardiovascular specialists in their referral to a surgeon – is the risk-adjusted mortality figure after bypass surgery in New York State. This figure is published annually on the Internet for all hospitals and surgeons providing such surgery in the state. It took, however, many years to develop this indicator and it is still hardly used for referral reasons. Cardiovascular specialists are of the opinion that the numbers are too hard to interpret since they are insufficiently valid and precise to them [17].

Looking at internal and external indicators from a theoretical point of view, now, strengthens the argument based on our empirical data. Waiting times – just as the above mentioned mortality rates – as they are published now can primarily be used for internal purposes because they only have meaning within a specific context. For outsiders it is difficult to understand the figures unless extensive validation is performed and supplementary information, for example on quality, is provided. In the next paragraph we will point at an initiative that makes waiting times relevant for specifically those actors that are supposed to act on them: patients and general practitioners.

4 Initiatives that make waiting time data relevant for patients and general practitioners

Waiting times, as we have seen above, can only be interpreted well within the context in which they are being generated. Patients, on the other hand, attach strong value to knowing how long they have to wait for their treatment. In addition, patients make use of several other considerations in choosing their hospital, e.g. the types of treatment that are offered, the organization of medical care and the experience of the doctor. Finally, patients trust upon their own experiences and the general practitioner's referral [11,18–20].

Offering actual waiting times, combined with information on (some of) these considerations, would then seem to be a big step forward in guiding patients in choosing a hospital. An example of a system that does just this, is a web-based application called 'CareDomain' (Zorg-Domein, www.zorgdomein.nl). This application is at this moment being

used by five hospitals and several hundreds of general practitioners in four regions in The Netherlands. Two other regions with five hospitals are at this moment implementing the application with approximately 250 general practitioners. This application was, originally, set up by a health insurance company in the end of the nineties in order to control the waiting times and improve the referral process between general practitioners and hospitals. For these goals, general practitioners and consulting physicians were brought together to agree on 'referral protocols', in which a variety of logistically inspired instruments were used. For example, general practitioners are offered different degrees of urgency when they refer patients: non-urgent, semi urgent, urgent. They can also choose between appointments with or without a reduced waiting time, depending on e.g. the patients' wish or severity of condition. In addition, general practitioners can choose for 'combination appointments', where patients can get a 'standard package' of investigations and appointments on 1 day. To facilitate this way of working, general practitioners and consulting physicians agreed on a set of standardized rules for diagnosing that were build into a decision support system. This decision support system was, at first, a paper-based system. But since this paper-based system was hard to update and a web-based application would make it possible to build in more functionalities, a software version was developed. This web page application allows the participating general practitioners to log on and to make electronic referrals to the hospital of choice on the basis of the above-mentioned agreements. Using this system makes it possible to refer most patients directly to the right outpatient clinic and consulting physician. In addition, this application allows general practitioners and consulting physicians to communicate and give feedback electronically. The core feature of this application is, however, that patients are referred on the basis of guaranteed waiting times that are connected to the different treatments in the different hospitals. This type of 'contract' is therefore completely different from the publication of waiting times on the Internet. Instead of giving waiting time data that are connected to an unknown context and therefore very hard to interpret for patients and general practitioners, waiting time becomes a parameter just as concrete as, e.g., the distance to the hospital. This means that waiting times actually become an element of choice. And since uncertainty in the date of treatment is seen as one of the most stressing elements by patients, this system fulfils a strong patient need [20].

In addition, the patient and the general practitioner can make a decision together regarding the choice of hospital [21]. Waiting times are in this way made interesting, relevant, valid and useful because there is a

connection between the waiting times itself and different other, relevant aspects of care. For the general practitioner, an important consideration is that the application is a result of an intensive cooperation process between consulting physicians and general practitioners. Indications for referral, for example, are the outcome of several meetings in which concrete agreements have been made regarding the goals for specific categories of patients. As a consequence, the general practitioners and the consulting physician know each other and each other's way of working better. Since patients especially seem to trust the quality of care perceived by the referring physician – instead of for example written down in public performance reports like report cards – this is a crucial element in helping the patient making health care decisions [19,21].

5 Discussion and conclusion

Waiting times, at first sight, seem to be a clear-cut process indicator of care. In the previous sections, however, we concluded that waiting times are ambiguous performance indicators that are difficult to interpret and serve different purposes for different actors. Waiting time data is infused with different meanings and connected to different interests depending on the actor's position in the system. Incentives, norms and traditions influence the generation and use of waiting time data. We also presented the example of an existing system that makes waiting time data available in a format that can be useful to several actors. In the following we will use these observations to (a) discuss the potentials of the publication of waiting time data as an instrument to help patients make health care decisions and (b) to elaborate on the relationships between patients and providers and between providers and third party payers.

Waiting time data may be useful to help patients choose providers. However, the results presented here indicate that general waiting time data are at best a crude indicator for individual patients that all have their own, specific needs. This is particularly true for conditions with an unclear diagnosis or conditions requiring multiple contacts with departments or consulting physicians that have different waiting times. Moreover, as stated earlier, studies also suggest that patients have complex preference structures when choosing hospitals, of which waiting time is only one element [11].

This raises several issues: are the costs of collecting waiting time data and the risk of misinformation too high to warrant such an effort? We will return to this question below. Another issue is: can we improve the usefulness of waiting time data for patients? As we stated in the

previous paragraph, in our view this is possible by offering patients actual (individually based) waiting times connected to other relevant data for the patient, like knowledge about treatment possibilities and the general practitioner's perceived quality of care of the consulting physician. In addition, since in the literature it is suggested that certainty of treatment time may be more relevant for patients than the actual wait [22,23], real time booking systems like 'ZorgDomein' seem to be the answer.

Are waiting time data useful indicators in the relationship between providers and third party payers? Again the answer must be that waiting time data at best provide a crude indicator, and as such may not be the most relevant indicator for creating incentive structures. The risk of rewarding the wrong providers for the wrong reasons is too high. In addition, it is well known from the literature that using performance indicators for this goal may result in perverse consequences such as gaming the numbers (e.g. favoring treatments where waiting times can be measured) [24] and might potentially lead to increased bureaucracy, less innovation and decreasing solidarity between hospitals [4,14,17,25-29]. All this leads us to conclude that waiting time data – when they are not related to other indicators – are a flawed indicator for creating incentives. Waiting time data should therefore primarily be used for reflective evaluations (self assessment) and as a starting point for dialogue, preferably in combination with other tools such as quality, service and patient satisfaction indicators. The information asymmetry and the risk of creating perverse incentives should be taken very seriously and the administrative costs of collecting and posting the data should be taken into account.

Are we then to abandon the use of waiting time indicators? No, but we must be very clear about the limitations, pitfalls and costs. In an ideal world the data needed to assess system performance has to do with optimal matching of scarce treatment resources and treatment needs evaluated by a complex set of medical factors, functionality, social and personal consequences, etc., rather than simple indicators such as waiting time. The main message of this study is to illustrate the complexity and the shortcomings of a PI like waiting time data. However, this is a message that may be difficult to sell in politicized health care systems that prefer easy and straightforward measurements. Waiting time data carries a huge potential for political symbolism. It is deceivingly easy to communicate, and it is difficult to argue against measures taken on the basis of waiting time data (who can be in favor of longer waits?). The use of indicators and accountability via measurement is also a central part of the New Public Management trend that has made an impact in most

health care systems. Waiting times have therefore attracted much attention in many countries. In some instances it has become a short hand for assessment of system performance and the many underlying parameters regarding quality, prioritization, service level, etc. This is unfortunate as it narrows the discussion and shifts the focus from complex relations to simple but insufficient and in some cases even harmful indicators.

References

1. Øvretveit J. Quality evaluation and indicator comparison in health care. International Journal of Health Planning and Management 2001;16:229–41.
2. Davies HTO, Hon MFPHM, Shields AV. Public trust and accountability for clinical importance: lessons from the national press reportage of the Bristol hearing. Journal of Evaluation in Clinical Practice 1999;5:335–42.
3. Marshall MN, Shekelle PG, Leatherman S, Brook RH. The public release of performance data. What do we expect to gain? A review of the evidence. Journal of the American Medical Association 2000;283:1866–74.
4. Øvretveit J. Informed choice? Health service quality and outcome information for patients. Health Policy 1996;37:75–90.
5. Appleby J, Harrison A, Dewar S. Patients choosing their hospital. British Medical Journal 2003;326:407–8.
6. Schauffler HH, Mordavsky JK. Consumer reports in health care: do they make a difference? Annual Review of Public Health 2001; 22:69–89.
7. Hibbard JH. Use of outcome data by purchasers and consumers: new strategies and new dilemmas. International Journal for Quality Improvement in Health Care 1998;10:503–8.
8. Stoop AP, Berg M, Juiste Tijd De. Totstandkoming Betekenis en Gebruik van wachttijdgegevens (The right time: realization, meaning and use of waiting time data). Den Haag: ZonMw; 2002.
9. School MAA. Het nut van wachttijdinformatie (The use of waiting time data). Utrecht: Prismant; 2001.
10. Brouwer WBF, Van Exel NJA, Stoop AP. Grensoverschrijdende zorg niet in trek (Little demand for cross-border care). Economisch Statistische Berichten 2002;87:30–1.
11. Brouwer W, Van Exel J, Hermans B, Stoop A. Should I stay or should I go? Waiting lists and cross-border care in the Netherlands. Health Policy 2003;63:289–98.

12. Hadjistavropoulos HD, Snider B, Bartlett G. Measuring the quality of performance in the management of waiting lists: using cataract surgery as an example. Journal of Quality Improvement 1998;24(8):407–22.
13. Light D. NHS waiting lists: the hidden agenda. Consumer Policy Review 2000;10:126–32.
14. Freeman T. Using performance indicators to improve heath care quality in the public sector: a review of the literature. Health Services Management Research 2002;15:123–37.
15. Thomson RG, Lally J. Performance management at the crossroads in the NHS: don't go into the red. Quality and Safety in Health Care 2000;9:201–2.
16. Berg M, Schellekens W. Paradigma's van kwaliteit De verschillen tussen externe en interne kwaliteitsindicatoren (Paradigms of quality: the differences between external and internal quality indicators). Medisch Contact 2002;57:1203–5.
17. Schneider EC, Epstein AM. Influence of cardiac-surgery performance reports on referral practices and access to care—a survey of cardiovascular specialists. New England Journal of Medicine 1996;335:251–6.
18. Lips JP. Ongrijpbare zorg (Intangible care). Medisch Contact 2002;57:1860–3.
19. Schneider EC, Epstein AM. Use of public performance reports: a survey of patients undergoing cardiac surgery. Journal of the American Medical Association 1999;279:1638–42.
20. Hadorn DC, Holmes AC. The New Zealand priority criteria project. Part 1: Overview. British Medical Journal 1997;314(7074):131–4.
21. Franken E. ZorgDomein en zijn domein effecten: een gebalanceerde effectevaluatie naar ZorgDomein (ZorgDomein and its domain-effects: a balanced effect-evaluation of ZorgDomein). Erasmus Medical Center; 2003.
22. Frankel SJ, West RR. Rationing and rationality in the National Health Service. In: The Persistence of Waiting Lists. London: MacMillan; 1993.
23. Brouwer WBF, Stoop AP. Vaste afspraken maken wachten draaglijker (Booked appointments make waiting more bearable). Medisch Contact 2001;56(23):908–9.
24. Davies HTO, Marshall MN. Public disclosure of performance data: does the public get what the public wants? The Lancet 1999;353: 1639–40.

25. Davies HTO, Lampel J. Trust in performance indicators? Quality in Health Care 1998;7:159–62.
26. Smith P. Measuring health system performance. European Journal of Health Economics 2002;3:145–8.
27. McColl A, Roderick P, Gabbay J, Smith H, Moore M. Performance indicators for primary care groups: an evidence based approach. British Medical Journal 1998;317(7169): 1354–60.
28. Pace N. Surgeons should beware of plans to print individual performance. British Medical Journal 2002;324:979.
29. Berg M. Informatie als instrument (Information as instrument). Rotterdam: iBMG, EUR; 2002.

Chapter 5

Stoop AP, Bal R, Berg M. OZIS and the politics of safety: Using ICT to create a regionally accessible patient medication record. *International Journal of Medical Informatics.* Accepted for publication.

Abstract

In studies on success and failure of ICT applications in health care, the 'context' is often used to explain the failure of a system and seldom to explain the success of a system. Science and Technology Studies (STS) have showed that for understanding success and failure of phenomena, one has to take a symmetrical approach and thus use the same concept for analyzing success and failure. In this study we analyze the success of OZIS, a communication protocol that makes it possible for pharmacists to exchange medication data by sharing a regionally accessible electronic medication record. Though OZIS serves a common goal – reducing medication errors – the stakeholders that are involved also have other, competing, interests. By focussing on the context and more specifically the interests of the stakeholders, we will show how the success of OZIS can be explained. By doing this, we will also show that this context is highly dynamic and that continuously changing incentives and constraints within the context lead to both facilitating *and* threatening the success of OZIS.

Keywords: Health Care, Information systems, Evaluation studies, Information Technology, STS

OZIS and the politics of safety: Using ICT to create a regionally accessible patient medication record

Arjen Stoop, Roland Bal, Marc Berg

1 Introduction

This study is about success and failure of ICT applications in health care, and the concepts that are used to explain success and failure. Interestingly, failures – if published upon – are both explained by referring to the imperfection of the system and by the context of use [1-5]. Successful (or potential successful) systems, however, are often only explained by referring to the 'intrinsic value' of the system or to the 'advantages' of the system compared to the 'old' situation without the system, see e.g. [6-8]. Success and failure are thus analyzed asymmetrically: the one by pointing at the intrinsic characteristics of the technology, the other by pointing at social or political mechanisms that prevent the technology to become a success. The effect of those analyses is that they create an opposition between social and technical types of explanation. Furthermore, they obfuscate the possibility that social and political aspects might have an explanatory value in success cases. Within Science and Technology Studies (STS), therefore, it has long been argued that for a thorough understanding of phenomena, one has to take a symmetrical approach and thus use the same concept for analyzing failure *and* success [9].

As 75 to 80% of ICT applications in health care are reported to fail, the explanation of success is of critical importance to medical informatics [10]. Success, however, is not much elaborated upon within the medical informatics literature. Explanations for failure, however, abound. Typically, explanations for failure are often sought in the mismatch between the technology and its context of use, sometimes in combination with wider social, organizational, economic or political contexts [1,2,11]. In this study, we will apply such contextual factors in trying to understand the success of an information technology, directing our intention to the economic and political context in which the technology is developed and used.

The technology we are focusing on is OZIS (Open Zorg Informatie Systeem, *Open Care Information System*), a communication protocol that makes it possible for pharmacists to exchange medication data by sharing a regionally accessible electronic medication record. In The Netherlands, patients normally get prescriptions for medication from their general practitioner or their medical specialist. Hence, they take their prescription to the pharmacist and pick up their medication. Though patients can go to any pharmacist they like, practically all patients only visit their 'own' pharmacist, which means that only that specific pharmacist has an (electronic) overview of the medication record of the patient [12]. This record mainly contains information on medication (prescribed and distributed medication, intolerances and contra-indications). In addition, one can find the patient's home address and type of insurance (sick-fund or private). However, when patients get a new prescription for medication during the nights and weekends, they have to go to the pharmacist that does shift by turn for the other (often small group of) pharmacists s/he works together with. More and more often, the patient has to go to the 'service pharmacist'. This 'service pharmacist' is a pharmacy that provides service only during the nights and the weekends and is responsible for the distribution of medication for a relatively large area. This means that patients are confronted with an unfamiliar pharmacy and vice versa. In both cases, there is the problem of getting access to the medication data of the patient.

Especially in these situations – where the pharmacist has no overview of the patient's medication – potentially dangerous situations can occur. Patients, for example, might already get certain medication that should not be taken at the same time as the newly prescribed medication. To prevent this 'break' in information exchange and to make it possible for pharmacies to have electronic access to the medication record of the patient, OZIS was developed.

The relevance of this case is clear: one of the major types of mistakes that are identified within health care are mistakes in the process of prescribing, transcribing, distribution and use of medication [13]. Despite differences in the use of definitions and methods, international publications show that about 0,3 – 3,6 % of all the hospitalized patients are confronted with one of these types of mistakes. Of all these mistakes, it is about 1% that experiences an Adverse Drug Event (ADE) [14]. Especially systems that make it possible to prevent 'breaks' in the process of prescribing, transcribing, distribution and use of medication between health care professionals, like OZIS, are seen as promising [15].

To make it possible for the pharmacist that works at nights or in the weekends to have access to the patients' medication data in that region by means of OZIS, it is necessary that as much pharmacists as possible are willing to work together and are connected to the OZIS infrastructure. However, working together means sharing patient data that normally are only accessible for the patient's 'own' pharmacist. This raises an interesting question: How can it be explained that Dutch pharmacists who are self-employed and therefore compete with each other, are willing to share their core business data and – by doing so – risk losing their customers (patients)? Perhaps even more important, the possibility of exchanging patient medication data between pharmacists that have a different pharmacist information system makes pharmacists at a sudden independent of this system (and vendor!). Vendors, potentially, open up the door for their customers (pharmacists) to contract one of the competitors. How can, in other words, the success of this ICT application – in terms of the willingness of pharmacists and vendors to work with OZIS – be explained in a competitive environment where the key stakeholders have exactly the same interest, i.e. maintaining or even maximizing their amount of customers?

We will, in the line of the theoretical standpoint described above, analyze the context in which OZIS has become a success in order to explain its success. Since the use of OZIS requires that stakeholders share their core business data – medication data of the patient – we decided to focus more specifically on the political and economic *interests* of the different stakeholders in the development, implementation and use of OZIS.

2 Methods

This report grew out of a study contracted by the Dutch Association of Pharmacists that wanted to know how many pharmacists used OZIS and whether they were satisfied. For this study, we decided to use qualitative as well as quantitative research methods. By contacting the vendors, we started with obtaining the addresses of all the pharmacies that (potentially) used the protocol. Subsequently, we conducted semi-structured interviews and observations of 12 pharmacies that used the OZIS protocol. These interviews and observations were directed at the experiences of the use of the system in daily practice and were meant to give input for a comprehensive questionnaire. Since OZIS is used in a regional setting, we were interested in the pharmacist's individual experiences in using OZIS as well as the experiences related to the collaboration between the pharmacists. We subsequently used this information to set

up a questionnaire that we sent to four pharmacists that were visited earlier. After minor revisions, we sent the questionnaire to all pharmacies that (possibly) used OZIS (n=485). In this questionnaire, we asked pharmacies to report their experiences on fixed categories. Practically all questions also offered the possibility of making comments in free text. We analyzed the data from the questionnaires with SPSSv12.

The result of this study was that OZIS was a (relative) success and we became intrigued with the question how this could be explained given the competitive environment in which the pharmacies worked. To answer this question, we interviewed eight key stakeholders: the designer of the OZIS protocol, three spokesmen of the different vendors that were and are directly involved in the implementation process of the protocol in the pharmacy practices in several regions in The Netherlands, the manager of the Dutch Association of Pharmacists, the chairman of the working group on electronic communication within the Dutch Association of Pharmacists, the chairman of the OZIS corporation and a pharmacist that is known for his innovative IT solutions and also is a user of OZIS. We conducted semi-structured interviews with these key persons and were specifically interested in their experiences with and perception of the evolution, implementation and success of the OZIS protocol. These interviews were literally transcribed and subsequently analyzed.

3 Results

The use of OZIS

The questionnaire was sent to 485 Dutch pharmacists that potentially used OZIS. 315 (65%) pharmacists returned the questionnaire. It appeared that 81% of the pharmacists used OZIS, while 15% was thinking about working with OZIS in the near future. The reasons for the pharmacists to work with OZIS appeared twofold: 85% of the pharmacists referred to the check on contra-indications and intolerances (like allergies) and 62% (also) referred to the possibility of checking the insurance status of the patient. This latter feature made billing much easier for pharmacists. Though the pharmacists thought that the system not always worked perfect and there was room for improvement, in general they were very satisfied. 82% of the pharmacists claimed that working with OZIS is necessary for the quality of pharmaceutical care. Though the number of pharmacists working with OZIS was strongly increasing[5], the questionnaires did not give insights in the exact reasons for

[5] In Amsterdam, for example, practically all pharmacists (about 100) started working with OZIS after we conducted our research.

pharmacists to start working with OZIS or the perception of the different stakeholders as to the reason of the success of OZIS. To find answers to these questions, we conducted eight additional interviews.

3.1 OZIS in its context: many stakeholders, many interests

The context in which OZIS was able to become a success consisted of several stakeholders. Four stakeholders were crucial: the Dutch Association of Pharmacists, the pharmacists themselves, the government and the vendors. The *common* interest of these stakeholders to build and implement OZIS started to exist around 1998 and resulted in the OZIS protocol that made it possible to exchange medication data electronically. Though it had – technically – been possible for many years to build OZIS, the above-mentioned stakeholders had too many competing interests. We will describe these competing interests and will show how a common interest was created and how OZIS developed from an idea to a working ICT application. Elaborating on this context shows how 'the technical' and 'the social' are intertwined and shows that the success of OZIS cannot be explained without paying attention to both dimensions. In addition, our analysis also shows how new incentives and constrains in the context in which OZIS is used, continuously enable *and* threaten the success of OZIS.

The Dutch Association of Pharmacists

The reason why The Dutch Association of Pharmacists was in favor of (a system like) OZIS was that it fulfilled the need of (a part of) the members of the association to have electronic access to the patient's medication data. Especially since pharmacists increasingly started working together during the nights and weekends in urban areas, access to patient medication data became crucial in order to guarantee high pharmaceutical care. Pharmacists were confronted with unknown patients and experienced the lack of access to medication data. More important, the association also had a growing political interest since OZIS would make it possible to fulfill the claim that pharmacists were *the* actor in health care that could live up to the responsibility of guarding the patient's medication safety. The association had already been claiming this for many years, but this claim was not taken serious since only patients' 'own' pharmacists had access to the patient's record.

> 'They (The Dutch Association of Pharmacists, AS) needed to live up to their claim that medication safety is the primate of the pharmacist. They had been claiming that for about twenty years, but could not fulfill this

> promise towards pharmacists, patients and politicians without pharmacists having access to the patient's record.' (Spokesman vendor A)

From the start of the 1990s the association was thinking about how to 'open up' the pharmacist information systems as to make it possible to share patient medication data. This idea was regularly communicated to the vendors, but nothing substantial happened. Consequently, at the end of 1995 the association installed a working group entitled: Werkgroep Geautomatiseerd Berichtenverkeer Apotheken, WGBA (*Working Group on Automated Message Reporting Pharmacies*). This working group was asked to investigate, among other things, the possibility of a communication protocol for electronic exchange of information. Though the working group concluded in 1996 that this protocol could be build relatively easily, the (OZIS) protocol, as we will see further on, was eventually build only in 2000.

The pharmacists

The pharmacists themselves could at that time (and still can) not be seen as a homogeneous group with clear and common interests. Pharmacists in urban areas were used to collaborate during the nights and weekends, while those in the countryside had always done their own shifts. While the former experienced the lack of access to patient medication data, the latter group did this to a lesser extent. In addition, while some regions were very steady regarding the amount of competition, other regions have experienced several initiatives of third parties – like general drugstores – that tried to take over a part of the market (see further). Especially in regions where there was an element of competition, there was (and sometimes still is) little trust in collaboration on the basis of exchange of patient medication data.

Despite these differences there was an increasing number of pharmacists that saw this lack of access to patient medication data as a problem. In 1995 there was a discussion in *Pharmaceutisch Weekblad* (Pharmaceutical weekly), the journal of the Dutch Association of Pharmacists, between a pharmacist and two vendors in which the pharmacist elaborated on the problems he experienced in his region in offering optimal medication safety during the nights and weekends, caused by the lack of access to the patients' records. In addition, the pharmacist claimed that patients were negatively surprised by the fact that pharmacists had no access to the records of patients that are clients to other pharmacists and that pharmacists' claim of medication safety therefore only appeared

to be a political claim, not one that was based on daily practice. These signals also reached the Dutch Association of Pharmacists.

> 'Many pharmacists wanted to have this access, but did not manage to realize this in their communication with the vendor. (Spokesman Dutch Association of Pharmacists).'

So the Dutch Association of Pharmacists as well as (a part of) the pharmacists experienced the lack of electronic access to patient medication data as a problem as early as the middle of the 1990s. Why then did nothing happen until five years later? Apart from the still heterogeneous interests of pharmacists, an important explanation for this lies with the vendors of pharmacist information systems.

The Vendors

In the Dutch healthcare market there were (and still are) three main vendors of pharmacist information systems. These vendors have always been competitors in a relatively small market and, as a consequence, reluctant to work together. This competition and reluctance to collaborate is again illustrated in *Pharmaceutisch Weekblad*. In 1995, the vendors discussed the possibility of using 'open standards' which would make communication possible between pharmacists with different pharmacist information systems. In this discussion, one vendor favored using open standards and also used one, another seemed willing but experienced technical difficulties, whereas the third refused to work with an open standard since working with open standards would make it impossible to guarantee privacy. Exchanging medication data, according to this vendor, should only be possible within 'formalized' forms of collaboration. By formalized groups this vendor referred to the 'clusters' in which pharmacists and GP's work together, using information systems from the same vendor. Sharing information outside these clusters was seen as 'irresponsible' according to the managing director of this vendor [16]. Because of their technical infrastructure, this vendor made it possible to create relatively large clusters with pharmacists. Within these clusters, pharmacists (and general practitioners) were able to exchange all kinds of patient data. Because of working with these clusters, new pharmacists (or pharmacists that used a different pharmacist information system) within those regions were forced to contract this vendor, if they wanted to work together with the other pharmacists. Clearly, the arguments of privacy and data quality this vendor was putting forward at the time, aligned strongly with its strive for market dominance in both the pharmacy and

GP sector. Economic interests were just as much at stake with the other vendors, the difference being that they did not focus on retaining or expanding their market share by creating their own - closed - clusters of pharmacies and GPs.

3.2 Money talks

The Dutch Association of Pharmacists saw competing interests between vendors as crucial for the delay in building the protocol by the vendors. Though the vendors worked together on the level of the design and implementation of standard modules, they were too much competitors to extend this cooperation into the use of open standards.

> 'Though it was possible to build the protocol, this didn't happen. First, this really required cooperation between the vendors, far beyond the cooperation that was needed for the implementation of, for example, a standard module for contra-indications. Second, this did not fit into the competitive relationship between the vendors.' (Spokesman Dutch Association of Pharmacists)

This deadlock, however, was broken in 1998 when the government - the fourth stakeholder - invested money in improving pharmaceutical care in The Netherlands by making agreements with the Dutch Association of Pharmacists (the so-called '*Meer Jaren Afspraken*' - Long Term Agreements). The Association decided to use this money to support the vendors financially in building and implementing the communication protocol in the pharmacist information systems. The vendors, subsequently, founded the OZIS corporation in 2000 with an independent chairman, builded the OZIS protocol and implemented it in the new releases of pharmacist information systems from that year on. From 2000 on, the number of pharmacists who electronically exchanged information during the nights and weekends by using OZIS was strongly increasing. This financial incentive, as the different stakeholders claim, was one of the crucial elements in understanding the success of OZIS.

> 'The real success of OZIS, in my opinion, was the moment that the Dutch Association of Pharmacists gave the vendors a bag of money. Showing the vendors that they were serious regarding the topic of information exchange, stimulated the vendors to move on this market.' (The designer of the OZIS protocol)

A second, more recent, incentive for collaboration between the vendors was the increasing role of communication in general.

> 'The vendors now realize that it is wise to standardize because that is where the money is. The more need for communication, the more work has to be done. This notion is present now, but was not there a few years ago.' (OZIS Chairman)

However, these were not the only incentives for the vendors to move on this part of the market. A third one came from legal developments in the European Community and the way these were enacted within the Netherlands. European legislation does not allow market parties to impede competition. And this was exactly what the vendor that worked with 'clusters' was accused of by one of the other vendors. Working in this cluster structure made it impossible for pharmacists outside the cluster (but within the region) to exchange information. The new legislation, however, forced this vendor to open up the system.

> 'Working together with (this vendor) was very hard. They have always had large regions with pharmacists connected to a cluster, without the need to communicate with other pharmacists. Until this new law, our pharmacists in those regions were not allowed to work together during the night and weekend shifts if they did not switch over [to this other vendor, AS].' (Spokesman vendor B)

Summarizing up till now, we have seen that until 1998 there was a deadlock that prevented vendors from collaborating in building a communication protocol that allowed to exchange patient data electronically, despite the fact that other stakeholders (patients, the Dutch Association of Pharmacists, individual pharmacists) started to develop common interests in the use of a standard like OZIS. Patients, first, wondered why pharmacists did not have access to their medication data and stimulated the pharmacists to discuss this possibility. Second, a growing part of the pharmacists wanted to have access to 'other' patients' data during the night and weekend shifts in order to guarantee the quality of pharmaceutical care (and facilitate billing). Third, the Dutch Association of Pharmacists wanted to live up to the promise that the pharmacist should be seen as *the* actor that can live up to patient safety regarding pharmaceutical care, this being one of the main arguments they used in the constant discussions with the Dutch government over the status of pharmaceutical care. It was only because of a clear incentive (the vendors being financed

via the government by the Dutch Association of Pharmacists) and a clear constraint (European legislation forcing the vendors' to open up their systems) that suddenly there was a growing – common – interest of the vendors in working together. It was this alignment of interests that made OZIS – in terms of the growing number of users, their satisfaction with OZIS and the perceived improvement of the quality of pharmaceutical care – into a success at the beginning of the 2000s.

Recently, however, several developments have threatened the alignment of interests that was so carefully built in the latter half of the 1990s. These developments also threaten the further actual implementation of OZIS. In the next paragraph we will point at an – from the pharmacists' perspective – unexpected threat.

3.3 New third parties on the pharmaceutical market and the use of OZIS

In 1999 the government changed the policy regarding the establishment of new pharmacies. Before that time, pharmacists were only allowed to work in a pharmacy that was owned by them. Since 1999, however, all kinds of enterprises like supermarkets are allowed to hire pharmacists (or sell a part of their building to a pharmacist) to sell prescription drugs. This change of the law was at the start not perceived as a serious threat to the pharmacists since very few initiatives were undertaken and practically all of these initiatives failed.

Recently however, all kinds of third parties, mainly (chains of) general drugstores, have hired pharmacists to sell prescription drugs. In addition, there is a growing tendency of hospitals to start outpatient clinic pharmacies. This means that patients, instead of leaving the hospital with a prescription, leave the hospital with their medication. Especially in the case of patients who are under supervision of a medical specialist and therefore visit the hospital regularly (and often use a relatively high amount of medication), the outpatient clinic pharmacy is a serious competitor for the community pharmacies. In some regions, new pharmacists have started outpatient clinic pharmacies; in other regions pharmacists who already have their own business in the region collectively establish outpatient clinic pharmacies and clear the profits of the medication prescribed by the medical specialist in the outpatient clinic. The former are seen as competitors, the latter not.

Though the situation that is described above is not completely new, the difference is that the government is now actively stimulating general drugstores in selling prescription drugs, as a part of its policy to cut on medication costs. This pressure on the pharmacists has even more

increased by recent reports that claim that the additional value of pharmacists, compared to for example general drugstores or even mail-order-firms, is lacking [17]. As a result, the Dutch Association of Pharmacists has been very intent to show the added value of the 'real pharmacist' compared to the general drugstore with a pharmacist working in it by pointing at the extra service provided to the patient. And by pointing at OZIS, the association can show this additional value in terms of the role of the pharmacist as *the* actor in health care that has a complete medication record of the patient.

However, the outpatient clinic pharmacies and general drugstores that have started to sell prescription drugs have indicated that they *also* wanted to make use of OZIS. This, of course, is a serious threat to the perceived *added* value of OZIS for the pharmacists compared to the general drugstores and competing outpatient clinic pharmacies. As a consequence, in several regions in the Netherlands pharmacists have refused to allow these third parties access to OZIS. The *Nederlandse Mededingingsautoriteit* (NMa, Netherlands Competition Authority) stated in a recent ruling that by European law pharmacists *have* to give general drugstores access to OZIS. It lead subsequently to several court cases in The Netherlands in the last few months in which pharmacies were forced to open up their systems for third parties in the region (e.g. Assen, Breda) [18,19].

This development has put pharmacists in a double bind: on the one hand they do not want to share the patient medication data with – in their view – illegitimate competitors. As they are forced to share OZIS with anyone entering the market, pharmacies are becoming reluctant to use OZIS for themselves. On the other hand they realize that OZIS lives up to their long lasting claim that they can guarantee the highest pharmaceutical care. From this perspective, they will have to continue (or begin) using OZIS. As the chair of the Dutch Association of Pharmacists who is also a pharmacist put it:

> 'In the current political climate of competition, OZIS would never have become a success. The pharmacists in Assen (a place in the northern part of The Netherlands where pharmacists were forced to share data with a general drugstore) would never have implemented OZIS if they knew they had to share their medication data.'

In practice however, the number of pharmacists implementing OZIS – despite the fact that third parties cannot be excluded from using

OZIS – is still increasing. This, according to a pharmacist, is very understandable:

> 'From a professional point of view regarding safety, pharmacists simply cannot afford to refrain from implementing OZIS.'

Depending on the nature of the regional competition, this respondent argued, the further implementation of OZIS could be easy, or not.

> 'The problem is that the government simultaneously wants to increase the competition within the market and cooperation between the stakeholders, which is contradictory. But to assure high quality of care the stakeholders have to work together. Depending on the regional context and contacts between pharmacists, this problem exists to a different extent.'

Though the pharmacists have been put in a double bind, they do not really have a choice: not using OZIS leads to sub optimal pharmaceutical care and is considered unprofessional behavior. In addition, using OZIS remains an opportunity to show added value – at least as long as not all general drugstores sell prescription drugs and demand the use of OZIS.

4 Conclusion

The success of OZIS, as we have seen in the above, is not about the technological quality of OZIS as a regionally accessible patient medication record, nor about the unambiguous context in which OZIS is used. The reason that the OZIS protocol after several years was built and implemented is about the *moment* in time in which the interests of all the relevant stakeholders converged: the Government wanted to see pharmaceutical care improved and allocated budget; the vendors were forced to the use of open standards by means of European law and were able to build OZIS with money from the Dutch Association of Pharmacists; the pharmacists finally could live up to their promise of becoming *the* actor in health care that could guarantee high pharmaceutical care.

However, the context in which OZIS is used has remained highly dynamic, showing the fragility of the interest alignment. The government, for example, made OZIS possible by breaking through the existing deadlock by allocating budget, whereas the same government is now threatening the success of OZIS by stimulating competition within the field. Pharmacists, as a consequence, are unable to show the added value

of the use of OZIS in their practice vis-à-vis third parties like general drugstores since third parties are now also allowed to make use of OZIS. In addition, the context also remains sub optimal from the perspective of the other involved stakeholders. Though the vendors realize that communication in general becomes more important and is good for their business, not all vendors are equally interested in opening up their system. European legislation, however, has forced the vendors to do this. From the pharmacists point of view OZIS offers the opportunity of improving their pharmaceutical care and legitimizing their position, but third parties' rights to make use of OZIS is clearly undesirable from their perspective.

The success of OZIS, hence, is not a given but should be seen as interplay between the stakeholders and their interests that has resulted in a highly dynamic setting with continuously changing incentives and constraints that has lead to both facilitating *and* threatening the success of OZIS. The reason that OZIS has survived up till now and probably also will survive in the future, is that certain steps have been made that make it almost impossible for the stakeholders to stop working with OZIS. That is, since OZIS is there and used by at least a part of the pharmacists, it has become connected to the very notion of what pharmaceutical care is all about.

Success and failure of information technology in health care, to conclude, is just as much about the 'intrinsic value' of the system, as about the social and political context in which the system is used. *Understanding* success and failure, therefore, can only be done by taking a symmetrical approach in which the same concepts are used for explaining success and failure.

References

1. Van't Riet A, Berg M, Hiddema F, Sol K. Meeting patients' needs with patient information systems: potential benefits of qualitative research methods. Intern J Med Inform 2001;64:1-14.
2. Stoop AP, Van't Riet A, Berg M. Using IT for Patient Education: Realizing surplus value? Patient Educ Couns 2004;54:187-195.
3. Forsythe DE. Blaming the user in medical informatics: The cultural nature of scientific practice. Knowledge and society 1992;9:95-111.
4. Forsythe DE. New Bottles, Old Wine: Hidden Cultural Assumptions in a Computerized Explanation System for Migraine Sufferers. Med Anthropol Quarterly 1996;10:551-74.
5. Forsythe DE, Brostoff MS, Buchanan BG. Information Exchange and the Problem of Perspective: What Should an Explanation System

Explain? Submitted to the 16th Symposium on Computer Applications in Medical Care 1993.
6. Buchanan BG, Moore JD, Forsythe DE, Carenini G, Ohlsson S, Banks G. An intelligent interactive system for delivering individualized information to patients. Artif Intell Med 1995;7(2):117-54.
7. Bobb A, Gleason K, Husch M, Feinglass J, Yarnold PR, Noskin GA. The Epidemiology of Prescribing Errors: The Potential Impact of Computerized Prescriber Order Entry. Arch Intern Med 2004; 164(7):785-792.
8. Buchanan BG, Carenini G, Mittal VO, Moore JD. Designing computer-based frameworks that facilitate doctor-patient collaboration. Artif Intell Med 1998;12(2):169-91.
9. Bijker WE, Hughes TP, Pinch T. The Social Construction of Technological Systems. New Directions in the Sociology and History of Technology. Cambridge, MA: MIT Press, 1987.
10. Berg M. The Search for Synergy: Interrelating medical work and patient care information systems. Meth Inf Med 2003;42:337-44.
11. Berg M, Aarts JECM, Van der Lei J. ICT in Health Care: Sociotechnical Approaches. Meth Inf Med 2003;42:297-301.
12. Van der Schee E, Van Dijk L, Blom L, Lelie-van der Zande R, Broerse A. Consumentenpanel Gezondheidszorg peilt leemtes (Consumers panel Health Care investigates ommissions. Pharmaceutisch Weekblad 2004;139:618-622.
13. Kohn LT, Corrigan JM, Donaldson MS, ed. To Err Is Human. Building a Safer Health System. Washington, D.C.: National Academy Press, 1999.
14. Van den Bemt PM, Egberts TC, De Jong-van den Berg LT, Brouwers JR. Drug related problems in hospitalized patients. Drug Safety 2000;22:321-33.
15. Berg M, Machielsen H, Huisman E. Bytes op recept? Medicatiesystemen in de specialistische zorg (Bytes on prescription? Medication systems in specialised care). Rotterdam: iBMG, 2002.
16. Nederlof RA. Hoe dichtbij is elektronisch berichtenverkeer tussen apothekers? (III) (How close is electronic data exchange among pharmacists?). Pharmaceutisch Weekblad 1995;130:883.
17. Bakker P. Het kan echt: betere zorg voor minder geld. Sneller Beter - de logistiek in de zorg. (It really is possible: better care with less money. Faster Better - the logistics in health care), 2004.
18. Dienstapotheek regio Assen. Nederlandse Mededingingsautoriteit, 2004.
19. Stichting Automatisering Gezondheidszorg Breda. Nederlandse Mededingingsautoriteit, 2004.

Conclusion

The studies in this thesis were related to the topic of evaluation of ICT applications in health care. The reason to focus on this topic was that though there had been a strong increase in the use of information systems in the last twenty years, the field of evaluation had received relatively little attention. Many systems were not evaluated and those that were, appeared not to be (that) successful. This situation, however, has changed to a certain extent. The field of evaluation has become visible, with discussions on the use of methods, the need for evaluation frameworks and guidelines and the potential additional value of computer-based systems versus paper-based systems. The aim of this thesis was to contribute to these discussions by addressing the following research questions:

1 In what way can evaluation studies benefit from the strengths and weaknesses of the different existing evaluation methods so they are better able to explain success and failure of information systems?
2 How can these insights be applied in real-life settings, what kind of possible constraints are observed and how can these constraints be dealt with?
3 Do computer-based information systems have the potential to realize surplus value compared to 'traditional' information systems used for education and information by patients and professionals and if not, how can this be explained?

These research questions were addressed in two parts. The first part of this thesis consisted of two chapters that gave an answer on the first two research questions by drawing on literature study and existing case material. The second part of this thesis consisted of three chapters in which the results from the first part were applied and in which the third research question was answered by means of four case studies.

The starting point of both parts was a sociotechnical perspective on success and failure of information systems. A sociotechnical perspective stresses the importance of the interrelation between technology and its social and political environment to understand success and failure. Interrelation refers to how 'the technology' (the information system) and 'the social' (the users and the social and political context they are part of) interact. From a sociotechnical perspective, success and failure of information systems are the *outcome* of this interaction.

I will first present the conclusions of the studies (chapters one to five) and finish with a general conclusion.

Chapter 1 discussed the perspective of the organizational decision-maker who is confronted with the need to acquire, implement and/or manage a Patient Care Information System (PCIS). Analysing the literature I concluded that the field of evaluation is scattered. The types of evaluation questions that are asked and the methods that can be used are infinite and badly demarcated. Different stakeholders, moreover, have different priorities in evaluating ICT. In this chapter, I argued that two dimensions of PCIS evaluation are crucial for evaluating ICT applications: the *domain* of evaluation and the different *phases* of PCIS implementation. Combining these two dimensions – by putting the domains on the x-axis and the phases on the y-axis – a table emerged in which each cell contained a distinctive set of evaluation questions. Regarding methods, I argued that the Randomised Controlled Trial (RCT) type of design is unsuitable for the organizational decision maker. RCTs are based on controlled, laboratory conditions, and are unsuitable for explaining why and how a PCIS is being used, or not. Integrating qualitative and quantitative methods by using results from one method as input for the other often offers the best way of answering evaluation questions. Which methods should be used depends on the type of questions and the required 'hardness' of the research data.

Chapter 2 took this analysis a step further by stressing the importance of evaluation as a balancing act between identifying specific and feasible evaluation questions, utilizing the amount of resources available and specifying the sufficient scientific 'hardness' of the data. In practice, I argued, there has to be a balance between the wish to generate robust data and the reasonableness of the effort to generate data. Experimental designs, for example, generate 'hard' data compared to non-experimental designs like case-control or before-after studies, but they also require much more effort. One can often generate sufficient data by doing a quick scan of the practices in which the system is used, for example through brief participant observations and interviews. In addition, I claimed that if the evaluation is meant to improve the implementation process and successful use of the system – which in practice practically always is desirable – it is crucial to apply a *formative* design instead of a *summative* design. Formative designs, in contrast to summative designs, are part of the implementation process and consequently have the potential to influence the course of the process for the better. To conduct an evaluation, a six-step model was proposed and analysed in terms of potential pitfalls and challenges in practice. The following steps were distinguished: agree why an evaluation is needed (step one), agree when

to evaluate (step two), agree what to evaluate (step three), agree how to evaluate (step four), analyse and report (step five) and assess and act (step six). The case studies illustrated why it is almost always necessary to do concessions towards the ideal situation. Limited resources, ever changing and multidimensional aims, and the changing environment in which any project is situated are the *de facto* situation – not an exception.

In chapters 3 to 5, evaluation studies of four information systems were presented. In chapter three and four the focus was on information systems to be used by patients and – to a lesser extent – general practitioners. Chapter five analysed an information system directed at professionals (community pharmacists).

In chapter 3, two patient information systems were evaluated using a formative design and qualitative and quantitative methods. One was to be used by patients in the waiting room of the GP, the other at home by children with amblyopia and their parents. Both systems were meant to replace already existing, traditional means of patient education such as brochures, leaflets and, to a certain extent, face-to-face communication. Both systems, however, were hardly used; and when they were used, the patients were predominantly dissatisfied. Importantly, using a formative design prevented the implementation of both patient information systems on a large scale, which was the initial starting point of the designers. In this chapter, I showed that the information systems failed because the emotional and cognitive content of the information was not geared towards the actual needs of the patients. Also, the moment in the illness course and the setting in which the information was offered, were problematic. In addition, I argued that building successful computer-based patient information systems is not only a complex task because of the reasons mentioned above. Information technologies, the Internet in general and health care IT applications specifically, are often mentioned as *the* solution for the limitations of 'traditional' means for patient education. As a consequence, the expectations of users are often very high. Building successful computer-based patient information systems, I concluded, requires systems that have rather significant additional advantages compared to traditional means of patient education.

In chapter 4, a formative approach using strictly qualitative methods was used to analyse the publication of waiting times on the Internet, meant to help consumers/patients and general practitioners act upon this information. Using information technology in this way is seen as

important from the perspective of the government, hospitals and third party payers since the public presentation of performance data is seen as a prerequisite for the development of a properly functioning health care *market*. Again, the results clearly stressed the importance of being aware of the tacit (and problematic) assumptions of the designers. Though waiting times, at first sight, seem to be clear-cut and straightforward 'bits' of information, it appeared that waiting times are ambiguous performance indicators that are difficult to interpret for those that are not aware of the way they are generated and the context in which they are generated. Publishing waiting times for *external* use, I concluded therefore, is problematic. Using waiting times for *internal* purposes like quality improvement and informing management, I argued however, is possible because the professionals and managers know the context in which they are generated. Waiting time data, however, *may* be useful to help patients choose providers. By offering patients *actual* waiting times *in addition to* other relevant data for the patient, such as information on treatment possibilities, the general practitioner's perceived quality of care of the consulting physician and guidance how to go to another hospital, the usefulness of publishing waiting time data could be increased.

Chapter 5, finally, was about using a software protocol (OZIS) for exchanging patient medication data between community pharmacists and the asymmetry in the concepts that are often used to explain success and failure of ICT applications. The success of OZIS, I concluded, cannot be explained by referring to 'traditional' reasons for success and failure. The reason that the OZIS protocol after several years was build and implemented was about the *moment* in time in which the interests of all the relevant stakeholders converged. This context, however, appeared highly dynamic. New – and unexpected – developments required that the stakeholders kept an eye on each other's behaviour and acted in their own advantage. In addition, the context remained suboptimal from the perspective of the involved stakeholders. The success of OZIS, hence, should be seen as interplay between the stakeholders and their interests that resulted in a highly dynamic setting with continuously changing incentives and constraints both facilitating *and* threatening the success of OZIS. The success of OZIS, I showed in this chapter, was just as much about the 'intrinsic value' of the system, as about the social and political context in which the system was used. Taking a symmetrical approach that uses the same concepts for analyzing success and failure, therefore, was a prerequisite for *understanding* the results of this study.

What can be concluded from the results mentioned above? First, I conclude that the field of evaluation of ICT applications in health care is changing. Whereas quantitative methods – with the RCT as golden standard – have been dominant within evaluation studies for many years, there is increasing attention for the use of qualitative methods or an integration of quantitative and qualitative methods [1]. Second, I conclude that there has been a shift in the topics that evaluation studies address. There is a shift of attention towards people, and organizational and social issues [2]. This shift has also proven to be *a condition* for understanding success and failure of ICT applications in health care. Research increasingly shows that success and failure of applications is about the *interaction* between the users and the technology, on several levels. Looking at the individual user – professional or patient – using a system, it is crucial that the designers are aware of their implicit and explicit assumptions of the use of the system. The case-studies in chapters three and four showed how a mismatch between the users' needs according to the designers and the users' needs in practice, is bound to lead to a failure of the system. In these cases, I do not suggest that the initial assumptions of designers were completely wrong, but they were too little geared to what patients really needed. Looking at the level of groups of professionals – in chapter five – it is clear that the joint use of systems requires an *organizational change*. After all, the existence of OZIS forced vendors and community pharmacists to make joint agreements on how to work together and even to take a position towards other actors, such as the general drugstores. The process in which a system becomes a success, in other words, is as much a technical as it is a social, organizational and a political process. This sociotechnical conclusion has proven to be essential in *understanding* how and why systems are used, or not [3]. The outcome of such a process can be hard to predict. Unexpected incentives during or after development or implementation of a system, may lead to more enthusiasm towards – or sabotage of – the system. Evaluation, my final conclusion, is necessary throughout the implementation process since the context is constantly changing which leads to shifting needs and, consequently, shifting incentives to use an information system, or not. Especially in clinical settings there often still is a demand for summative evaluation methods, which in practice appears to be highly problematic [4,5]. A formative approach towards evaluation as (partly) applied in the chapters three to five means that the evaluation results become part of the implementation process itself. Applying a formative approach in which evaluation not only is part of the implementation process but also of the design process, offers even

more possibilities. Though the problematic assumptions of the systems that are evaluated in chapter three were – as mentioned – not that wrong, they have never been subject to evaluation in the design phase. If these systems had been evaluated by using a formative approach starting the evaluation in the design phase, some of these assumptions probably would have been problematized, which could have lead to more successful systems or refraining from building the systems at all.

References

1. Ammenwerth E, De Keizer N. An Inventory of Evaluation Studies of Information Technology in Health Care. Methods of Information in Medicine 2005;44:44-56.
2. Kaplan B, Shaw NT. Future Directions in Evaluation Research: People, Organizational, and Social Issues. Meth Inf Med 2004;43: 215-31.
3. Berg M. Patient care information systems and health care work: a sociotechnical approach. Intern J Med Inform 1999;55:87-101.
4. May C, Ellis NT. When protocols fail: technical evaluation, biomedical knowledge, and the social production of 'facts' about a telemedicine clinic. Social Science & Medicine 2001;53:989-1002.
5. De Bont A, Bal R, De Mul M, Berg M. Face to face in a database: On trust and distrust, task delegation and professionals. Social Science & Medicine 2005; Submitted for publication.

Samenvatting

Dit proefschrift levert een bijdrage aan discussies rondom het nut en de noodzaak van evaluatie van ICT applicaties in de gezondheidszorg. Onder ICT applicaties worden in dit proefschrift zowel op computertechnologie gebaseerde informatiesystemen voor artsen, verpleegkundigen en andere gezondheidszorgprofessionals verstaan als informatiesystemen die (primair) voor patiënten ontworpen zijn. De eerstgenoemde groep informatiesystemen worden ook wel Patiënten Zorg Informatie Systemen (PZIS) genoemd en zijn systemen die informatie verwerken die binnen het primaire proces van patiëntenzorg is gegenereerd en/of hierbinnen wordt gebruikt, zoals elektronische patiënten dossiers. Informatiesystemen voor patiënten verwijzen naar die systemen die primair voor gebruik door patiënten zijn ontwikkeld, zoals systemen waarin medische informatie kan worden opgezocht.

De relevantie van dit onderwerp is gelegen in het gegeven dat vele implementatietrajecten van informatiesystemen mislukken doordat behoeftenonderzoek bij gebruikers ontbreekt, er sprake is van een gebrekkige implementatie of dat het informatiesysteem slecht aansluit bij het dagelijkse werk van zorgprofessionals. Er wordt in dit proefschrift ingegaan op de relatie tussen het ontwerpen, implementeren en gebruiken van bovengenoemde soorten systemen en de rol die evaluatie kan spelen bij het verhogen van de succeskansen hierbij. Hierbij is een sociotechnisch perspectief gehanteerd. Dit perspectief behelst een visie waarin het belang van de wisselwerking tussen technologie en de bestaande praktische, sociale en politieke omgeving wordt benadrukt om succes en falen van technologieën te begrijpen. Vanuit dit perspectief is het slagen of mislukken van een technologie de *uitkomst* van deze wisselwerking.

De onderzoeksvragen waren als volgt:

1 Op welke wijze kunnen evaluatiestudies gebruik maken van de sterke en zwakke punten van de verschillende bestaande evaluatiemethoden zodat succes en falen van informatiesystemen beter kunnen worden verklaard?
2 Hoe kunnen deze inzichten worden toegepast in de praktijk, welke mogelijke beperkingen worden waargenomen en hoe kan het beste met deze beperkingen worden omgegaan?
3 Hebben op computertechnologie gebaseerde informatiesystemen potentiële meerwaarde ten opzichte van 'traditionele' informatiesystemen ten behoeve van voorlichting aan patiënten en gezondheidszorgprofessionals en zo ja, waaruit bestaat deze meerwaarde?

Ik gebruik in dit onderzoek vooral, maar niet uitsluitend, kwalitatieve onderzoeksmethoden als documentanalyse, observaties en interviews.

Allereerst wordt in hoofdstuk 1 ingegaan op de problematiek van de veelheid van onderzoeksvragen en –methoden die potentieel relevant kunnen zijn bij de evaluatie van een informatiesysteem. Deze problematiek kan worden teruggedrongen door het implementatietraject van een informatiesysteem in twee dimensies op te delen – het *domein* van evaluatie en de *fase* waarin de implementatie zich bevindt. Deze categorisering geeft richting aan de keuzes die kunnen worden gemaakt. Door de twee genoemde dimensies in een kruistabel te zetten, ontstaat een nog specifieker overzicht van evaluatievragen die relevant kunnen zijn. Afhankelijk van de vraag kunnen zowel verschillende kwalitatieve als kwantitatieve methoden worden gebruikt. Door onderzoeksresultaten van kwalitatieve en kwantitatieve methoden te *integreren*, echter, kunnen de rijkste onderzoeksresultaten worden verkregen.

Hoofdstuk 2 geeft een visie op het in de praktijk uitvoeren van evaluatieonderzoek in een zestal stappen. De belangrijkste conclusie van dit hoofdstuk is dat in de praktijk altijd een balans gezocht zal moeten worden tussen het identificeren van specifieke en haalbare evaluatievragen, het optimaal gebruiken van de beschikbare middelen (tijd, geld, deskundigheid) en het specificeren van de benodigde objectiviteit van de te verkrijgen gegevens. Door snel maar weloverwogen te werk te gaan en de waarde van verkregen onderzoeksresultaten niet aan dominante wetenschappelijke paradigma's te toetsen maar aan de vraag die op dat moment antwoord vereist, kan zeer waardevolle informatie worden verkregen met relatief beperkte middelen.

Hoofdstuk 3 behandelt de vraag in hoeverre op computertechnologie gebaseerde informatiesystemen meerwaarde hebben ten opzichte van traditionele middelen voor patiëntenvoorlichting als video, brochures, folders en – tot op zekere hoogte – direct patiëntencontact. Bestudering van literatuur en case-studies laat zien dat zowel de meer traditionele middelen voor patiëntenvoorlichting als de nieuwe op computertechnologie gebaseerde informatiesystemen ieder zijn sterke en zwakke kanten kennen. Laatstgenoemde informatiesystemen dienen dan ook primair als een aanvullende vorm van informatievoorziening te worden gezien en niet – waar voorstanders voor pleiten – als vervanging van traditionele vormen van patiëntenvoorlichting. Het vervangen van papieren patiëntenfolders door een pc met medische informatie voor patiënten in de wachtkamer van de huisarts, schiet dan ook zijn doel voorbij.

Hoofdstuk 4 gaat in op de (on)mogelijkheid van het inzetten van informatiesystemen voor het informeren van patiënten en huisartsen over de lengte van de wachttijden voor behandeling in het ziekenhuis. Het louter publiceren van wachttijden, zo luidt de conclusie in dit hoofdstuk, zal er niet toe leiden dat patiënten en huisartsen zonder meer kiezen voor het ziekenhuis met de kortste wachttijd. Daarvoor is een prestatieindicator als wachttijd te complex. Zo is het voor de patiënt (noch de huisarts) relevant te weten hoe lang de wachttijd voor een hernia-operatie is als niet duidelijk is om wat voor type operatie het gaat en welke specialist deze uitvoert. Bovendien spelen in het keuzeproces voor een ziekenhuis tal van andere aspecten een rol, zoals fysieke nabijheid. Informatiesystemen waarbij ook andere – voor de patiënt belangrijke – informatie wordt verstrekt bieden wel perspectief.

Het laatste hoofdstuk gaat in op het succes van een informatiesysteem voor openbare apothekers (OZIS) dat als doel heeft medicatiefouten terug te dringen door toegang te bieden tot een regionaal elektronisch patiënten medicatiedossier. In dit hoofdstuk wordt specifiek gekeken naar de rol van de context waarbinnen OZIS een succes is geworden. Nadere analyse van de context laat zien dat deze zeer dynamisch is en hierdoor tevens fragiel. Echter, doordat het gebruik van OZIS onderdeel is gaan uitmaken van professionele farmaceutische zorg lijkt de verdere implementatie – ondanks toegenomen tegenstrijdige belangen van de betrokken actoren – niet meer tegen te houden. Waar 'de context' veelal wordt gebruikt voor het verklaren van falende implementatietrajecten van informatiesystemen, illustreert dit hoofdstuk het belang van een symmetrische benadering van succes en falen van informatiesystemen.

Wat kan uit bovenstaande worden geconcludeerd? Ten eerste concludeer ik dat er binnen het vakgebied van evaluatie van ICT-applicaties in de zorg de afgelopen tien jaar meerdere ontwikkelingen hebben plaatsgevonden. Zo is er toenemende aandacht voor de beperkingen van kwantitatieve onderzoeksmethoden en de (potentiële) meerwaarde van een gecombineerd gebruik van kwantitatieve en kwalitatieve onderzoeksmethoden. Deze combinaties worden in de praktijk echter nog weinig toegepast. Ten tweede concludeer ik dat er sprake is van een verschuiving van aandacht van technische aspecten naar de rol van menselijke, sociale en organisatorische aspecten. Deze verschuiving blijkt een *voorwaarde* te zijn voor het begrijpen van succes en falen van informatiesystemen in de zorg. Ten derde concludeer ik dat het noodzakelijk is reeds tijdens het implementatieproces te evalueren. Het louter achteraf evalueren van een informatiesysteem levert niet voldoende informatie op om een verklaring

te kunnen geven voor het succes of falen van een informatiesysteem. Nog belangrijker, evaluatie gedurende de implementatie-fase kan bijdragen aan het succes van de implementatie zélf.

Curriculum Vitae

Arjen Stoop was born April 13, 1971 in Echten (Friesland), The Netherlands. In 1988 he graduated from the Drachtster Lyceum and studied nursing at the Rijkshogeschool Groningen from 1988 to 1992. From 1992 to 1995 he studied health sciences at Maastricht University. Subsequently he has been working at the Netherlands Institute for Health Services Research (NIVEL) and at the Foundation for Harmonization of Quality Review in Health Care and Welfare (Stichting HKZ), both in Utrecht. Since 1999 he has been working at the Institute of Health Policy and Management of the Erasmus MC in Rotterdam as a research fellow. He developed a particular interest in evaluation of information systems in health care and the use of ICT for quality improvement. From 2004 on he has – within his function as a research fellow – partly been working as a consultant in *Sneller Beter*, a national project for quality improvement in Dutch hospitals. He is married to Karen Ingversen. They have one daughter, Thera Sofia, who is almost two years old.