ASTHMACRITIC

COMPUTER-BASED CRITIQUING IN DAILY PRACTICE

MANON M. KUILBOER

AsthmaCritic

Computer-based critiquing in daily practice

Acknowledgements

The research published in this thesis was financially supported by The Netherlands Asthma Foundation (#92.62).

This thesis has been published with the financial support from

- The Netherlands Asthma Foundation
- Microbais Automatisering BV

Cover design and thesis layout:	Karin P. Kuilboer, Catching, Amersfoort
Clothes styling:	Renate M. Kuilboer, Amersfoort
Printed by:	Optima Grafische Communicatie, Rotterdam

Kuilboer, M.M.
AsthmaCritic; Assessment of the feasibility and effect of computer-based critiquing on asthma and COPD management in daily practice
Thesis Erasmus University Rotterdam – with summary in Dutch

ISBN: 90-6734-087-1

AsthmaCritic

Assessment of the feasibility and effect of computer-based critiquing on asthma and COPD management in daily practice

AsthmaCritic

Het vaststellen van de haalbaarheid en effect van het bekritiseren van het medisch handelen rond astma en COPD door een computerprogramma in de dagelijkse praktijk

Proefschrift

ter verkrijging van de graad van doctor aan de
Erasmus Universiteit Rotterdam
op gezag van de
Rector Magnificus

Prof.dr.ir. J.H. van Bemmel

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

woensdag 8 januari **2003 om 15:45 uur**
door
Manon Marguerite Kuilboer

geboren te Amsterdam

Promotiecommissie

Promotor:	Prof. dr. J. van der Lei
Overige leden:	Prof. dr. A. Hasman
	Prof. dr. S. Thomas
	Prof. dr. Th. Stijnen
Copromotor:	Dr. M.A.M. van Wijk

Aan Eric, Titia, Otto en Sascha

Aan pappa, mamma, Renate en Karin

Contents

1

INTRODUCTION

BACKGROUND

THE KNOWLEDGE-PERFORMANCE GAP

Medical knowledge is changing rapidly[1]. Physicians have difficulty staying up to date with the changes; it can take many years for new knowledge to be integrated into daily practice. Delays in the integration of new knowledge can lead to suboptimal care and unnecessary health-care expenditures[2-5]. Development of techniques that bring physician behavior more in line with current knowledge is the goal of active research[6, 7]. Experience with existing techniques has shown that timing is important for effective recall and application of recommended practices[8]. Passive techniques that ignore timing and that do not provide information at the moment the physician needs it most have been shown to have minimal effect on a physician's knowledge and behaviour[9, 10]. The increased use of computers in daily practice creates an opportunity to use computerized decision-support systems (CDSSs) to introduce new medical knowledge precisely at physicians' moments of interest – during daily practice.

COMPUTERIZED DECISION-SUPPORT SYSTEMS

Studies have shown that CDSSs can change physician behavior[11-13]. In these studies, the success of systems in daily practice, was shown to be highly dependent on the extent to which the systems had been integrated into a physician's workflow. This observation led to a renewed emphasis on the need to integrate CDSSs with clinical information systems[14]. In the Netherlands, over 80-90% of the general practitioners use an electronic patient record[15]. Physicians record their patient data in the electronic patient record themselves, during the patient encounter. This high percentage of general practitioners using an electronic patient record instead of a paper-based record makes general practice a suitable environment in which to evaluate the feasibility of a CDSS integrated with physicians' electronic patient records. If the CDSS is integrated with an information system that the general practitioner is already accustomed to using, then we can study the feasibility of support generated as a by-product of general practitioners' information systems[16].

GUIDELINES

Professional health-care organizations develop clinical guidelines to help physicians treat patients according to the latest medical knowledge[17, 18]. Practice guidelines summarize large volumes of clinical evidence, and provide practical recommendations that are tailored to daily practice. Guidelines, like other passive techniques, however, have not been effective in changing physician behavior[19-21]. In

addition, the number of guidelines issued has become so great that physicians cannot manage them all in daily practice[22]. However, the content of the guidelines is extremely valuable because guideline authors have successfully assimilated current evidence and other information such as policies, preferences, and resource availability. Thus, text-based guidelines can provide a good starting point for the development of a CDSS[8]. In the Netherlands, the Dutch College of General Practitioners has been publishing practice guidelines since 1989. Because the Dutch practice guidelines are issued by an authoritative organization, and are well accepted by practitioners, they provide a resource that we can use to develop a trustworthy knowledge base[23]. In our research, we will use the Dutch guidelines as the starting point for our CDSS[18].

CRITIQUING SYSTEMS

One concern with widespread use of CDSSs is that physicians will run the risk of becoming dependent on CDSSs if decisions are made for them. They may become passive in their decision making and, therefore, more vulnerable to mistakes[24, 25]. To overcome this problem, a type of CDSS that does not make decisions for the user has emerged. A system of this type critiques decisions the physician has already made. Such systems, called critiquing systems, generate feedback based on a physician's treatment plan. The feedback is based on information recorded in the electronic patient record[16]. When physicians use critiquing systems, they continue to make their own decisions, and those decisions are subsequently evaluated by the software. This process is called the critiquing model[26]. Relatively few critiquing systems have been developed since they were first introduced in the early eighties[27] One reason for the infrequent development of critiquing systems is that the critiquing dialogue is complex, and in the past, information systems that could be used as a data source were rare. Therefore, the timing of feedback could not be optimized[26]. One system that did succeed in the early days by providing feedback at the time of patient care was the HELP hospital information system. The HELP system successfully generated reminders to physicians as a byproduct of patient data recording activities[28]. In our study, we further explore the feasibility and effect of the critiquing-system approach. We have designed and developed a system that critiques treatment plans of general practitioners, and we implemented the system by integrating it with an electronic patient record. The system provides decision support to physicians in their daily practice.

DOMAIN

Since it is not (yet) realistic to develop a system that can cover all of medical practice, we limited ourselves to the development of a CDSS in one domain. The choice of the domain was based first, on the rate of recent changes in recommendations for diagnosis and treatment, and second, on the proportion of the Dutch population that may be affected. Diagnosis and treatment of asthma and chronic obstructive pulmonary disease (COPD) have changed considerably in recent years; and the short time intervals between consecutive publications of guidelines for asthma and COPD[29-35] demonstrate this rapid change. In addition, studies have shown that the treatment of asthma and COPD lags behind current recommendations published in clinical guidelines, and results in unnecessary high health-care expenditures[3, 36-38]. Since up to about a third of a population suffers from asthma or COPD-related symptoms, a significant gain in health-care quality may be achieved if the care that is provided is consistent with current guidelines[39-43]. We, therefore, choose asthma and COPD as the domain in which to evaluate our ideas.

Summarizing, in this thesis we try to answer the following question:

WHAT IS THE FEASIBILITY AND EFFECT OF A CRITIQUING SYSTEM INTEGRATED WITH AN ELECTRONIC PATIENT RECORD IN GENERAL PRACTITIONERS' DAILY PRACTICE IN THE DOMAIN OF ASTHMA AND COPD?

THESIS OVERVIEW

This thesis encompasses four steps in the software-development process; *simulation, implementation*, *testing,* and *evaluation*. The chapters in this thesis follow these four steps.

SIMULATION

In the Netherlands, most general practitioners use an electronic patient record to record their patient data. The amount of data and information recorded may be sufficient to fulfill the needs of a practicing physician, but may be insufficient to fulfill a critiquer's needs. Typically, the physician needs enough data to serve as a reminder of past events in order to provide care for a patient on subsequent visits. In contrast, a CDSS needs specific data elements that may or may not be recorded in order to draw conclusions. In addition, if the system needs data that are missing and prompts the

physician to enter that data, the physician may find it annoying to be disturbed by interruptions. Time is limited in the practice setting, and interruptions may not be appreciated, even if they come from a supportive instrument. Therefore, we had to ask the question:

"DO GENERAL PRACTITIONERS RECORD ENOUGH DATA ELECTRONICALLY SUCH THAT CRITIQUING CAN BE PROVIDED BASED ON THESE DATA ONLY?"

To answer this question, we present in Chapter 2 the results of a simulation study in which four reviewers, playing the role of the computer, generated critiquing comments on electronic medical records of patients with asthma or COPD. Three general practitioners, playing the role of the users, assessed these comments and provided missing information when requested. The reviewers reevaluated their critiquing comments after the missing information had been provided. The results of this study gave insight into the feasibility of using electronic patient records as the single data source for a critiquing system, and it addressed the question of whether it was necessary for the system to ask the physician for missing data.

IMPLEMENTATION

In the literature, there is no blueprint available for critiquing systems that serve the needs of general practice. Therefore, we asked the following question:

"WHAT ARE THE REQUIREMENTS FOR A NON-INQUISITIVE CRITIQUING SYSTEM THAT CRITIQUES THE CARE PROVIDED BY GENERAL PRACTITIONERS?"

To answer this question, we describe and discuss in Chapter 3 the functional design and implementation of a non-inquisitive critiquing system called AsthmaCritic. We acknowledge that different users may want to have different levels of control, and recognize that we do not know much about which characteristics of CDSSs determine a good fit between a CDSS and a working environment.

TESTING

After we completed the simulation study to test the feasibility of our ideas, and after we built the system, we needed to test the quality of the system's critiques before the system could be put into practice. The question, therefore, was:

"CAN A CDSS PERFORM THE ROLE OF HUMAN REVIEWERS IN THE DOMAIN OF ASTHMA AND COPD?"

To answer this question, we evaluate in Chapter 4 the performance of AsthmaCritic in a laboratory setting. The question was if critiquing in the domain of asthma and COPD would work with a critiquing system instead of with human reviewers. To address this question, we let the system analyze over 100.000 electronic patient records and assessed its performance. In doing so, we also assessed the system's robustness – that is, we assessed whether it functioned reliably. We could not install a system that regularly shows unexpected functioning into a physician's clinical practice. In the discussion in Chapter 4, we reflect on the number and kind of comments generated by the system with respect to physician responsibility in decision-making.

EVALUATION

The final step in a software development process is the evaluation of the object of interest in its intended working environment. In our evaluation, we were interested in how effective the system was in general practice. We, therefore, asked the following question:

"IS ASTHMACRITIC ABLE TO CHANGE GENERAL PRACTITIONERS' MONITORING AND TREATMENT OF PATIENTS WITH ASTHMA AND COPD?"

To answer this question, we describe in Chapter 5 the results of a randomized controlled trial with AsthmaCritic in daily practice. We describe the effect of the non-inquisitive critiquing system on general practitioners' monitoring and treatment of their patients with asthma and COPD. We discuss the meaning and limitations of our results.

SUMMARY AND FUTURE RESEARCH

We conclude with Chapter 6 in which we summarize this work and make suggestions for future research.

APPENDIX

The Appendix (CD-ROM) contains a demo of AsthmaCritic, its manual, and the description of its knowledge base.

REFERENCES

1. Wyatt J. Uses and sources of medical knowledge. Lancet 1991;338:1368-72.
2. Sudlow M, Rodgers H, Kenny RA, Thomson R. Population based study of use of anticoagulants among patients with atrial fibrillation in the community. Bmj 1997;314(7093):1529-30.
3. Stoloff S. Current asthma management: the performance gap and economic consequences. Am J Manag Care 2000;6(17 Suppl):S918-25; discussion S925-9.
4. Amsterdam EA, Laslett L, Diercks D, Kirk JD. Reducing the knowledge-practice gap in the management of patients with cardiovascular disease. Prev Cardiol 2002;5(1):12-5.
5. Haines A, Jones R. Implementing findings of research [see comments]. British Medical Journal 1994;308(6942):1488-92.
6. Smith WR. Evidence for the effectiveness of techniques To change physician behavior. Chest 2000;118(2 Suppl):8S-17S.
7. Grol R. Personal paper. Beliefs and evidence in changing clinical practice. Bmj 1997;315(7105):418-21.
8. Wyatt JC. Clinical knowledge and practice in the information age: a handbook for health professionals. London: The Royal Society for Medicine Press limited; 2001.
9. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance. A systematic review of the effect of continuing medical education strategies. Jama 1995;274(9):700-5.
10. Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. The Cochrane Effective Practice and Organization of Care Review Group. Bmj 1998;317(7156):465-8.
11. Hunt DL, MD, Haynes RB, MD, PhD, Hanna SE, MA, PhD, Smith K. Effects of computer-based clinical decision support systems on physician performance and patient outcomes: A systematic review. Journal of the American Medical Association 1998;280:1339-46.
12. Mitchell E, Sullivan F. A descriptive feast but an evaluative famine: systematic review of published articles on primary care computing during 1980-97. British Medical Journal 2001;322(7281):279-82.
13. van Wijk MA, van der Lei J, Mosseveld M, Bohnen AM, van Bemmel JH. Assessment of decision support for blood test ordering in primary care. a randomized trial. Ann Intern Med 2001;134(4):274-81.
14. Lobach DF, MD PhD, Underwood HR, MD MBA. Computer-based decision support systems for implementing clinical practice guidelines. Drug Benefit Trends 1998;10(10):48-53.
15. Wolters E. Evaluatie invoering electronisch voorschrijfsysteem. Monitoring: de situatie in 2002. Utrecht: NIVEL; 2002.

16. Lei van der J, Musen MA. A model for critiquing based on automated medical records. Computers and Biomedical Research 1991;24:344-78.

17. Guidelines on the management of asthma. Statement by the British Thoracic Society, the Brit. Paediatric Association, the Research Unit of the Royal College of Physicians of London, the King's Fund Centre, the National Asthma Campaign, the Royal College of General Practitioners, the General Practitioners in Asthma Group, the Brit. Assoc. of Accident and Emergency Medicine, and the Brit. Paediatric Respiratory Group. Thorax 1993;48(2 Suppl):S1-24.

18. NHG. 'NHG Standaarden' (Dutch College of General Practitioners Guidelines). In. 2002 ed: NHG; 2002.

19. Davis DA, Taylor-Vaisey A. Translating guidelines into practice. A systematic review of theoretic concepts, practical experience and research evidence in the adoption of clinical practice guidelines. Cmaj 1997;157(4):408-16.

20. Feder G, Eccles M, Grol R, Griffiths C, Grimshaw J. Clinical guidelines: using clinical guidelines. Bmj 1999;318(7185):728-30.

21. Cabana MD, Rand CS, Powe NR, Wu AW, Wilson MH, Abboud PA, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. Jama 1999;282(15):1458-65.

22. Hibble A, Kanka D, Pencheon D, Pooles F. Guidelines in general practice: the new Tower of Babel? Bmj 1998;317(7162):862-3.

23. Grol R, Thomas S, Roberts R. Development and implementation of guidelines for family practice: lessons from The Netherlands [editorial]. J Fam Pract 1995;40(5):435-9.

24. Denig P, Haaijer-Ruskamp FM. Therapeutic decision making of physicians. Pharm Weekbl Sci 1992;14(1):9-15.

25. Morris AH. Computerized protocols and bedside decision support. Crit Care Clin 1999;15(3):523-45, vi.

26. Shortliffe EH. Computer programs to support clinical decision making. Jama 1987;258(1):61-6.

27. Miller PL. A critiquing approach to expert computer advice: ATTENDING. Boston: Pittman; 1984.

28. Evans RS, Larsen RA, Burke JP, Gardner RM, Meier FA, Jacobson JA, et al. Computer surveillance of hospital-acquired infections and antibiotic use. Jama 1986;256(8):1007-11.

29. Bottema BJAM, Fabels EJ, Van Grunsven PM, Van Hensbergen W, Muris JWM, Van Schayck CP, et al. NHG Standaard CARA bij Volwassenen: Diagnostiek [Guidelines of the Dutch College of General Practitioners: Chronic Respiratory Diseases in Adults: Diagnostics]. Huisarts en Wetenschap 1992;35(11):430-6.

30. Dirksen WJ, Geyer RMM, De Haan M, Kolnaar BGM, Merkx JAM, Romeijnders ACM, et al. NHG Standaard Astma bij Kinderen [Guidelines of the Dutch College of General Practitioners: Asthma in Children]. Huisarts en Wetenschap 1992;35(9):355-62.

31. Waart van der MAC, Dekker FW, Nijhoff S, Thiadens HA, Van Weel C, Helder M, et al. NHG Standaard CARA bij Volwassenen: Behandeling. [Guidelines of the Dutch College of General Practitioners: Chronic

Respiratory Diseases in Adults: Therapy]. Huisarts en Wetenschap 1992;35(11):437-43.
32. Dirksen WJ, Geijer RMM, De Haan M, De Koning G, Flikweert S, Kolnaar BGM. NHG-standaard astma bij kinderen [Guidelines of the Dutch College of General Practitioners: Asthma in Children]. Huisarts en wetenschap 1998;41(3):130-43.
33. Geijer RMM, Thiadens HA, Smeele IJM, Zwan van der AAC, Sachs APE, Bottema BJAM, et al. NHG-Standaard COPD en astma bij volwassenen: Diagnostiek [Guidelines of the Dutch College of General Practitioners: Chronic Obstructive Respiratory Diseases and Asthma in Adults: Diagnostics]. Huisarts en Wetenschap 1997;40(9):415-28.
34. Geijer RMM, Hensbergen van W, Bottema BJAM, Schayck van CP, Sachs APE, Smeele IJM, et al. NHG-Standaard astma bij volwassenen: Behandeling [Guidelines of the Dutch College of General Practitioners: Asthma in Adults: Therapy]. Huisarts en wetenschap 1997;40(9):443-54.
35. Geijer RMM, Schayck van CP, Weel van C, Sachs APE, Zwan van der AAC, Bottema BJAM, et al. NHG-Standaard COPD: Behandeling [Guidelines of the Dutch College of General Practitioners: Chronic Obstructive Respiratory Diseases: Therapy]. Huisarts en wetenschap 1997;40(9):430-42.
36. Smeele IJ, Van Schayck CP, Van Den Bosch WJ, Van Den Hoogen HJ, Muris JW, Grol RP. [Discrepancy between the guidelines and practice by family physicians in treating adults with an exacerbation of asthma or chronic obstructive pulmonary disease]. Ned Tijdschr Geneeskd 1998;142(42):2304-8.
37. Rabe KF, Vermeire PA, Soriano JB, Maier WC. Clinical management of asthma in 1999: the Asthma Insights and Reality in Europe (AIRE) study. Eur Respir J 2000;16(5):802-7.
38. Taylor DM, Auble TE, Calhoun WJ, Mosesso VN, Jr. Current outpatient management of asthma shows poor compliance with International Consensus Guidelines. Chest 1999;116(6):1638-45.
39. Worldwide variations in the prevalence of asthma symptoms: the International Study of Asthma and Allergies in Childhood (ISAAC). Eur Respir J 1998;12(2):315-35.
40. Variations in the prevalence of respiratory symptoms, self-reported asthma attacks, and use of asthma medication in the European Community Respiratory Health Survey (ECRHS). Eur Respir J 1996;9(4):687-95.
41. Mannino DM. COPD: epidemiology, prevalence, morbidity and mortality, and disease heterogeneity. Chest 2002;121(5 Suppl):121S-126S.
42. Yawn BP, Wollan P, Kurland M, Scanlon P. A longitudinal study of the prevalence of asthma in a community population of school-age children. J Pediatr 2002;140(5):576-81.
43. Tirimanna PR, van Schayck CP, den Otter JJ, van Weel C, van Herwaarden CL, van den Boom G, et al. Prevalence of asthma and COPD in general practice in 1992: has it changed since 1977? [see comments]. Br J Gen Pract 1996;46(406):277-81.

2

SIMULATING AN INTEGRATED CRITIQUING SYSTEM

Published in the Journal of the American Medical Informatics Association; 1998; 5: 194-202

Manon M. Kuilboer
Johan van der Lei
Johan C. de Jongste
Shelley E. Overbeek
Ben Ponsioen
Jan H. van Bemmel

ABSTRACT

Objective: To investigate factors that determine the feasibility and effectiveness of a critiquing system for asthma/COPD that will be integrated with a general practitioner's (GP's) information system.

Design: A simulation study. Four reviewers, playing the role of the computer, generated critiquing comments and requests for additional information on six electronic medical records of patients with asthma/COPD. Three GPs who treated the patients, playing users, assessed the comments and provided missing information when requested. The GPs were asked why requested missing information was unavailable and why requested missing information that was available had not been recorded. The reviewers reevaluated their comments after receiving requested missing information.

Measurements: Descriptions of the number and nature of critiquing comments and requests for missing information. Assessment by the GPs of the critiquing comments in terms of agreement with each comment and judgment of its relevance, both on a five-point scale. Analysis of causes for the (un)availability of requested missing information. Assessment of the impact of missing information on the generation of critiquing comments.

Results: Four reviewers provided 74 different critiquing comments on 87 visits in six electronic medical records. Most were about prescriptions (N=28) and the GPs' workplans (N=27). The GPs valued comments about diagnostics the most. The correlation between the GP's agreement and relevance scores was 0.65. However, the GPs' agreements with prescription comments (complete disagreement, 31.3%; disagreement, 20.0%; neutral, 13.8%; agreement, 17.5%; complete agreement, 17.5%) differed from their judgments of these comments' relevance (completely irrelevant, 9.0%; irrelevant, 24.4%; neutral, 24.4%; relevant, 32.1%; completely relevant, 10.3%). The GPs were able to provide answers to 64% of the 90 requests for missing information. Reasons available information had not been recorded were: the GPs had not recorded the information explicitly; they had assumed it to be common knowledge; it was available elsewhere in the record. Reasons information was unavailable were: the decision had been made by another; the GP had not

recorded the information at the time of the encounter. The reviewers left 74% of the comments unchanged after receiving requested missing information.

Conclusion: Human reviewers can generate comments based on information currently available in electronic medical records of patients with asthma/COPD. The GPs valued comments regarding the diagnostic process the most. Although they judged prescription comments relevant, they often strongly disagreed with them, a discrepancy that poses a challenge for the presentation of critiquing comments for the future critiquing system. Requested additional information that was provided by the GPs, led to few changes. Therefore, as system developers, faced with the decision to build an integrated, *non-inquisitive* or an *inquisitive* critiquing system, the authors choose the former.

INTRODUCTION

Decision-support systems have shown to be able to provide users with support[1-3]. Most of these systems, however, have failed to get incorporated into daily clinical practice[4, 5]. The main reason for this failure is the failure to meet the specific requirements of the future users, resulting in a mismatch between problem and solution[6]. For example, the system requires special data entry which interferes with normal practice, it is too time consuming for daily use, the system's timing does not fit the clinical routine, or it ignores the physician's intelligence[7-9].

Researchers have argued that decision-support systems need to be integrated with electronic medical records to improve these systems' chances to be incorporated into the physician's daily routine[7, 10]. Such an integration with the electronic medical record allows a decision-support system to review or critique the physician's treatment using the data already available in the electronic medical record. In The Netherlands, over 50% of the general practitioners have been using an electronic medical record for several years, making the time ripe for the development of integrated decision-support systems[10]. We are developing a particular kind of integrated decision-support systems, *critiquing systems,* that generate critiquing comments based on the user's actions as recorded in these medical records[11-13].

Integrated critiquing systems aim to support physicians based on facts already entered in the electronic medical record, thus avoiding the problem of double data entry[4]. We are building integrated systems that will not ask the physician for additional data: *non-inquisitive* critiquing systems. The downside of this approach is the limited availability of data[14, 15]. That is, the ability of such a system to critique diagnosis and treatment is limited by the data available in the electronic medical record. If the electronic medical records do not contain sufficient data, the concept of an integrated, non-inquisitive critiquing system is unfeasible. To determine the feasibility of such a system, we need insight into the number and the nature of comments that can be made based upon the information in the electronic medical record.

If the lack of patient data in the record prohibits the development of a non-inquisitive critiquing system, we can consider a separate module that requests additional information. To determine the viability of a separate data collection module, we need insight into the availability of information missed from the record for the critiquing task. Such a module would be useful only when physicians are able to provide the required

information. In addition, we have to gain insight into the relevance of this information. When the impact of additional information on the generation of comments is small, obtaining the additional data may require too much effort on the part of the general practitioner.

Whether a critiquing system will be rejected or accepted is also determined by the users' judgment of its critiquing comments. To determine which critiquing comments might be perceived as inappropriate, builders of a critiquing program need insight into general practitioners' responses to these critiques[12].

Before building an integrated non-inquisitive critiquing system, a system builder thus has to face a number of questions, that center around two issues:

- Will it be possible to generate critiquing comments based on the information available in the electronic medical record, and how will general practitioners judge them?

- How much information is missing? Can general practitioners provide the missing information? Why and why not? Does provided information make a difference for the generation of critiquing comments?

In the past, we addressed these issues by building and evaluating prototypes[16]. This process, however, is very time-consuming. An alternative to building prototypes is to perform a simulation study in which humans play the role of the system. To our surprise, we have not found examples of studies using such an approach. The closest comparable technique is used in the field of human-computer interface (HCI). It is called the “Wizard-of-Oz” technique; to reveal important aspects of an interface design, humans play the role of a computer[17]. The user's commands are interpreted by humans, who, invisible to the users, generate the expected responses. The difference of our approach from the Wizard-of-Oz technique is that we do not blind our users for the fact that humans play the role of the computer system.

In this article, we report the results of a small-scale simulation study that attempted to answer the system builders' questions with regard to a critiquing system that supports general practitioners in the diagnosis and treatment of patients with asthma/chronic obstructive pulmonary disease (COPD).

METHODS

In this simulation study, we reviewed six medical records of patients who had been diagnosed as having chronic respiratory disease (asthma/COPD) by their general practitioners. The records were randomly selected from the electronic medical record systems of three general practitioners. In The Netherlands, most general practitioners make use of electronic medical records that adhere to the national standard prescribing the data elements that an electronic medical record should contain (WCIA)[18]. For our study, we worked with physicians who were using the general practitioners' information system ELIAS©, one of the most commonly used information systems for general practitioners in The Netherlands[10].

The role of the computer system was played by four reviewers with special interest in asthma/COPD: two specialists (one pulmonologist and one pediatric pulmonologist) and two general practitioners. The role of "users" was played by the same three general practitioners who provided the medical records. The simulation was conducted in three phases as illustrated in Figure 1.

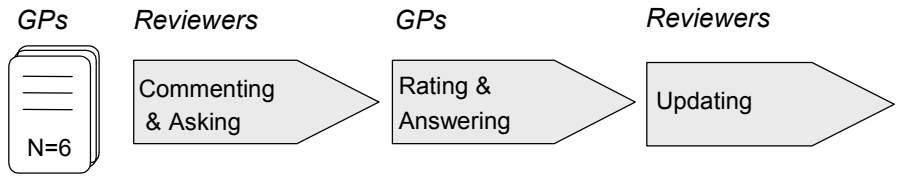


FIGURE 1.

FOUR REVIEWERS ANALYZED SIX MEDICAL RECORDS. THE REVIEWERS GENERATED COMMENTS AND REQUESTED FURTHER INFORMATION WHEN NEEDED. THE GENERAL PRACTITIONERS RATED THESE COMMENTS AND PROVIDED THE MISSING INFORMATION. WHEN INFORMATION WAS NOT AVAILABLE, THEY WERE ASKED TO EXPLAIN WHY. FINALLY, THE REVIEWERS UPDATED THEIR COMMENTS, TAKING THE ADDITIONAL INFORMATION INTO ACCOUNT.

REVIEWERS' COMMENTS AND REQUESTS FOR FURTHER INFORMATION

In the first phase of the study (see Figure 1), we provided each reviewer with the medical records. For each visit documented in the record, we asked the reviewer to formulate suggestions for changes in the physician's patient management – *critiquing comments*. Also, we asked the reviewer to verify whether the record contained sufficient information to comprehend the general practitioner's interventions. If the

reviewer felt that information was missing, we asked him to formulate this as a *request for additional information*.
As each reviewer worked independently, they sometimes used different formulations of essentially the same comment. To enable comparison, we mapped those comments to a single comment. Subsequently, we asked the reviewers to verify the mapping. Finally, we submitted all comments to all reviewers, and we asked each reviewer to indicate for each comment whether he agreed with it.

For the analysis, we assigned each comment to one of four categories:

- *Diagnostic* comments dealt with the diagnostic part of the doctor-patient encounter (examples: “Before the diagnosis asthma can be established, the presence of allergies should be investigated” and ”The child has an upper respiratory tract infection; she should have her ear, nose, and throat examined.”).
- *Workplan* comments dealt with the physician's proposed therapeutic strategy (examples: “The patient is using too many bronchodilating agents; anti-inflammatory therapy is indicated” and ”The child is taking ketotifen which is not indicated for children older than 4 years without frequent symptoms”).
- *Prescription* comments dealt with prescription specifications (examples: “The prescription frequency is too high” and ”The prescription of different routes of administration is irrational”).
- *Follow-up* comments dealt with the timing of a follow-up (examples: “The patient should return in six weeks instead of three months because his condition is instable” and ”The follow-up is insufficient because the effect of this nasal corticosteroid should be checked”).

Because the reviewers worked independently, different reviewers could also request identical additional information using slightly different wording. We mapped these requests from more than one reviewer to a single request. For the analysis, we assigned the requests for additional information to one of three categories. Two of the three categories dealt with missing facts, and one category dealt with missing reasoning:

- Requests about *Factual patient data* dealt with missing data of the medical history, physical examination, diagnosis, or additional tests (for example, “What did the pulmonary examination reveal?”, ”What are the patient's

symptoms after this period of two years?", " What is the patient's condition after treatment with inhalation corticosteroids?").

- Requests about *Factual therapeutic data* dealt with the physician's therapeutic interventions (for example, "What was the exact amount of medication?", "Which medication has been continued?", "How much corticosteroids has the patient been instructed to take per dosage?").
- Requests about *Motivation* dealt with missing information about the general practitioners' motivation for their policy (for example, "Why did the physician change the medical device?", "What was the indication for oxazepam – nocturnal asthma?", "Why doesn't the doctor do anything?").

General Practitioners' Ratings and Answers

In the second phase of the study, we asked the general practitioners to consider each individual comment and to rate its correctness (on a five-point scale ranging from complete disagreement to complete agreement) and its relevance (on a five-point scale ranging from completely irrelevant to completely relevant). The relevance of a comment was defined as "being relevant for this situation". In addition, we asked the general practitioners to answer the reviewers' requests for additional information. This question could result in one of two situations: 1) the physician could not provide the requested information, in which case he was asked to explain why; 2) if he could provide the requested information, he was asked to explain why he had not recorded the information in the first place.

Reassessment of the Comments by the Reviewers

In the third phase, we asked the reviewers to reassess their initial comments. We provided the reviewers with the original records, their comments, their requests for additional information, and the additional information given by the general practitioners. We subsequently gave the reviewers the opportunity to retain, withdraw, or change comments, or to add new comments.

Analysis

For analysis, we counted the comments per category and the requests for additional information per category. As an indication of agreement among the reviewers, we counted per comment the number of reviewers that agreed with that comment. To explore the comments' relevance and correctness as given by the general practitioners, we used descriptive analysis and calculated the correlation coefficient.

To analyze the causes for requested information to be (un-)available, we counted the reasons given by the general practitioners per category. To analyze the impact of additional information, we counted the number of comments that the reviewers left unchanged, withdrew, changed, or added.

RESULTS

The six patient records covered 87 visits, on average 14.5 (range: 5-24) visits per record. The reviewers made a total of 74 different comments, on average 0.9 per visit.

REVIEWERS' COMMENTS AND REQUESTS FOR FURTHER INFORMATION

CATEGORIES OF COMMENTS MADE BY THE REVIEWERS

The number of reviewers' comments per category is shown in Table 1. The largest categories of comments were related to *Prescriptions* (N=28; 38%) and the physician's *Workplan* (N=27; 36%).

CATEGORY	FREQUENCY	PERCENTAGE
Diagnostics	13	18%
Workplan	27	36%
Prescription	28	38%
Follow-up	6	8%
Total	74	100%

TABLE 1.

FREQUENCIES AND PERCENTAGES OF COMMENTS MADE BY REVIEWERS PER CATEGORY.

MISSING INFORMATION

The reviewers stated a total of 132 requests for additional information, which we mapped to 90 single requests. The percentage of each category of requests for additional information is shown in Figure 2.

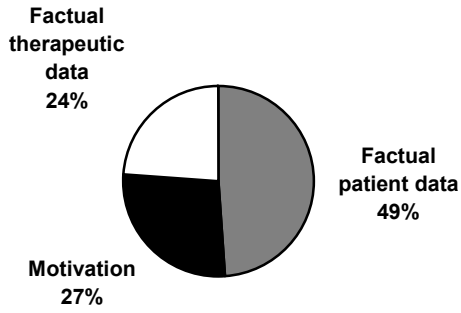


FIGURE 2.

SUMMARY OF INFORMATION MISSED IN SIX ELECTRONIC MEDICAL RECORDS BY REVIEWERS. THREE CATEGORIES OF MISSING INFORMATION COULD BE IDENTIFIED; *FACTUAL PATIENT DATA (N=44)* – ANY REQUEST FOR ADDITIONAL INFORMATION RELATED TO A PATIENT'S MEDICAL HISTORY, PHYSICAL EXAMINATION, DIAGNOSIS, OR ADDITIONAL TEST; *FACTUAL THERAPEUTIC DATA (N=22)* – REQUESTS ASKING THE PHYSICIAN ABOUT HIS OR HER THERAPEUTIC STRATEGY– ; *MOTIVATION (N=24)* – REQUESTS ASKING FOR THE PHYSICIAN'S MOTIVATION FOR HIS OR HER INTERVENTION–.

ASSESSMENT OF AGREEMENT AMONG THE REVIEWERS

Out of the 74 comments made by the reviewers, 45% were endorsed by all four reviewers, 31% by three, 12% by two, and 12% by only one expert. In two of the 74 comments, the reviewer who had stated the comment subsequently disagreed with his own comment.

GENERAL PRACTITIONERS' RATINGS OF COMMENTS

Each of the three general practitioners rated each individual comment for correctness and relevance on a five-point scale, resulting in 222 judgments of correctness and 222 judgments of relevance. Of these judgments, the general practitioners explicitly had no opinion in 9 (correctness) and 11 (relevance) cases. These judgments were excluded from further analysis.

Figure 3 shows the overall distribution of the general practitioners' judgments. The correlation coefficient between the three general practitioners' agreement scores and their relevance scores was r=0.65. The most frequently assigned scores were *agreement* (code: +1) and *relevant* (code: +1). In almost 20% of the cases the

general practitioners *completely disagreed* with a comment, but only 10% of the comments were judged *completely irrelevant*.

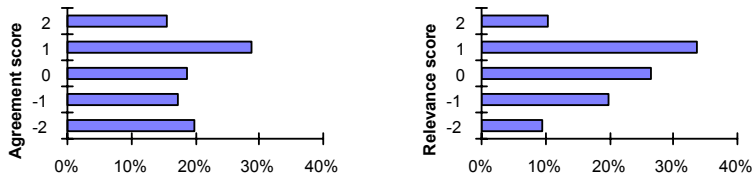


FIGURE 3.

DISTRIBUTION OF THE INDIVIDUAL AGREEMENT SCORES AND RELEVANCE SCORES (*N* SCORES = 424) OF THREE GENERAL PRACTITIONERS FOR COMMENTS (*N* COMMENTS =74) GENERATED BY REVIEWERS. THE VERTICAL AXES SHOWS THE RANGE OF THE SCORES THAT THE GENERAL PRACTITIONERS COULD ASSIGN (-2 REPRESENTING *COMPLETE DISAGREEMENT*, TO +2 REPRESENTING *COMPLETE AGREEMENT* AND -2 REPRESENTING *COMPLETELY IRRELEVANT*, TO +2 REPRESENTING *COMPLETELY RELEVANT*, RESPECTIVELY). THE HORIZONTAL AXES SHOW THE PERCENTAGES WITH WHICH EACH SCORE WAS ASSIGNED.

Figures 4 and 5 show the general practitioners' judgments for the four individual categories of comments *Diagnostics*, *Workplan*, *Prescription*, and *Follow-up*. Overall, the general practitioners rated the category of comments regarding *Diagnostics* positively, both for their agreement with a comment as well as their judgment of its relevance.

The agreement scores and relevance scores of the comments regarding the general practitioners' *Workplan* were also generally positive, even though 14% (11/80) of these judgments were *complete disagreement* and an equal percentage (also 11/80) were judged *completely irrelevant*. The general practitioners gave a relatively large number of comments in the category *Prescriptions* a negative agreement score (*complete disagreement*: 31% (25/80)). In contrast, the general practitioners were less negative about the relevance of these comments (*completely irrelevant*: 9% (7/78)).

GENERAL PRACTITIONERS' ANSWERS TO THE REQUESTS FOR FURTHER INFORMATION

The reviewers had stated 90 different requests for additional information. The general practitioners were able to provide information responding to 58 (64%) of the reviewers' requests. The reasons the information had not been recorded

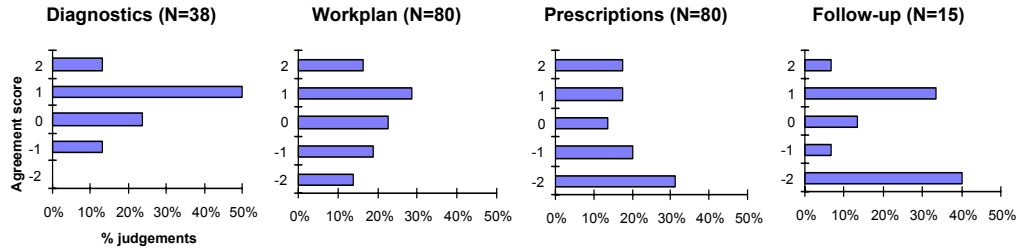


FIGURE 4.

AGREEMENT SCORES OF GENERAL PRACTITIONERS (N=213) FOR COMMENTS (*N* = 74) GENERATED BY REVIEWERS. THE RESULTS ARE SHOWN BY THE FOUR CATEGORIES OF COMMENTS; *DIAGNOSTICS* (*N* = 13), *WORKPLAN* (*N* = 27), *PRESCRIPTION* (*N* = 28), AND *FOLLOW-UP* (*N* = 6). FOR EACH CATEGORY, THE DISTRIBUTION OF THE AGREEMENT SCORES IS SHOWN BY THE HORIZONTAL BARS. THE VERTICAL AXES SHOW THE RANGE OF THE SCORES THAT THE GENERAL PRACTITIONERS COULD ASSIGN (-2 REPRESENTING *COMPLETE DISAGREEMENT* TO +2 REPRESENTING *COMPLETE AGREEMENT*). THE HORIZONTAL AXES SHOW THE FREQUENCIES WITH WHICH THE SCORES WERE GIVEN.

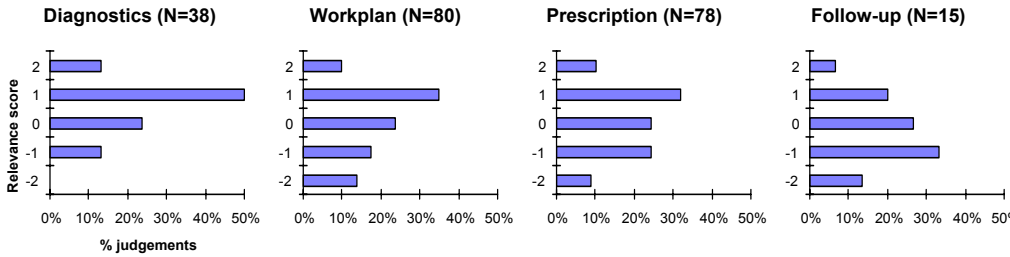


FIGURE 5.

RELEVANCE SCORES OF GENERAL PRACTITIONERS (N = 211) FOR COMMENTS (*N* = 74) MADE BY REVIEWERS. THE RESULTS ARE SHOWN BY THE FOUR CATEGORIES OF COMMENTS; *DIAGNOSTICS* (*N* = 13), *WORKPLAN* (*N* = 27), *PRESCRIPTION* (*N* = 28), AND *FOLLOW-UP* (*N* = 6). FOR EACH CATEGORY, THE DISTRIBUTION OF THE RELEVANCE SCORES IS SHOWN BY THE HORIZONTAL BARS. THE VERTICAL AXES SHOW THE RANGES OF THE SCORES THAT THE GENERAL PRACTITIONERS COULD ASSIGN (-2 REPRESENTING *COMPLETELY IRRELEVANT* TO +2 REPRESENTING *COMPLETELY RELEVANT*). THE HORIZONTAL AXES SHOW THE FREQUENCIES WITH WHICH THE SCORES WERE GIVEN.

in the medical record are summarized in Table 2. In 54% of the 58 answered requests, the physician indicated that the requested information had not been explicitly recorded in the medical record (e.g., why something had not been done). In 22%, the requested information had been assumed to be known (e.g., “fever” means a temperature above 38.5 degrees Celcius). In 17% of the cases, the requested information had been recorded elsewhere in the electronic medical record (e.g., information recorded as a personal note to the record). In 5% of the cases, the information had not been recorded in the electronic medical record yet, but had been available in the paper-based record. (In The Netherlands, most general practitioners use electronic medical records, while in the past, they used paper-based records. During the transition from paper-based records to electronic medical records, the two types of records temporally co-exist until all relevant medical data have been recorded electronically). Finally, in 2%, the information was provided by an external individual (e.g., a family member).

In 32 of the 90 requests (36%), the general practitioners were not able to provide the requested information. The reasons requested information was unavailable are summarized in Table 3. In 41% of these 32 cases the general practitioner indicated that the decision had been made by another individual (most commonly the general practitioner on call during the night or on weekends). In 37%, the physician did not know the answer to the request, nor did he know where to locate the missing information (e.g., information about the physical examination had not been recorded at the time of the visit).

FREQUENCY		**REASON**
NO.	**%**	
31/58	54%	Not explicitly recorded
13/58	22%	Assumed to be known
10/58	17%	Registered elsewhere in the electronic medical record
3/58	5%	Registered in the paper-based record
1/58	2%	Other source

TABLE 2.

REASONS INFORMATION THAT WAS AVAILABLE WHEN REQUESTED (*N* = 58), HAD NOT BEEN RECORDED IN THE MEDICAL RECORD.

In 19% of the cases, the physician knew where to find the information, but had not taken the effort to retrieve it (e.g., in the paper-based record). In 3%, the request could not be answered because it was unclear to the physician.

FREQUENCY		REASON
NO.	%	
13/32	41	Other decision maker
12/32	37	Information not known
6/32	19	Too much effort required
1/32	3	Request unclear

TABLE 3.

REASONS REQUESTED INFORMATION WAS UNAVAILABLE (*N*=32)

REASSESSMENT OF THE COMMENTS BY THE REVIEWERS

After the general practitioners had provided the requested additional information, the reviewers received the medical records, their comments, their requests for additional information, and the provided additional information to review their comments. The reviewers left 55 (74%) comments unchanged, withdrew 11 of them (15%), changed 8 (11%) comments, and made 15 new ones.

DISCUSSION

We performed a simulation study to gain insight into some of the issues that determine the feasibility and effectiveness of a computer-based critiquing system that will support general practitioners in the treatment of their patients with chronic lung diseases. In this simulation study, we focused on issues that center around the availability of medical data for critiquing, and the role of missing information. In addition, we investigated the kinds of comments that could be made and the general practitioners' assessments of these comments.

Our study scope was small and thus the potential for an extensive analysis was limited. A more extensive design would have made a more extensive analysis possible, but it would have cost more time and effort; the physicians need time and patience to work through the medical data, comments, and changes.

However, a more extensive study would have made it possible to analyze generated comments in relationship to the general practitioners' assessments on a more detailed level, making more detailed recommendations possible. In addition, instead of being purely descriptive, the analysis could have been extended to a statistical analysis of changes in comments. In retrospect, an extension of the study to include the general practitioners' reassessments of the edited comments would have been valuable.

The design of a simulation study depends on the lessons that need to be learned from it. An advantage of a simulation study of a computer system is that feedback is possible on issues that, when prototyping, could have emerged only at a very late stage. For example, the role of additional information could have been investigated only when additional modules had been programmed, and sufficient functionality would have been available for which this information would have made a difference. From our study, we could draw the conclusions that we required to determine some of the core aspects of our system design.

Critiquing systems that are integrated with a general practitioner’s information system work with medical data as they are currently available in general practitioners' electronic medical records. Therefore, the available data for the system will be limited to the data that a general practitioner is able and willing to enter into the electronic medical record. This is a potential limitation that may impair the generation of useful critiquing statements in clinical domains such as chronic lung diseases. However, as P. Miller pointed out, it remains to be seen whether a limited availability of data necessarily limits the effectiveness of a critiquing system[19]. In other words, there may be reasons why it is good to be generic. The acceptability of a system may improve when comments are less specific, because, for example, comments are less likely to be wrong.

As system developers, we are faced with the decision between a *non-inquisitive* and an *inquisitive* system design. Our simulation study showed that it is possible for human reviewers to generate critiquing comments (on average, one comment per visit), despite the fact that the reviewers in our study often missed information (90 requests for further information were stated over 87 visits). The majority of the comments (74%) were left unchanged by the reviewers after the participating general

practitioners had provided additional information for 64% of the requests. On the other hand, 26% of the comments were changed and 15 new comments were made, showing that additional information may change some comments or give rise to additional ones.

To assess the feasibility of an *inquisitive* design, we assessed the availability of missing information. In our study, one-third of the requests could not be answered. To explain why requested information was unavailable, the general practitioners most often mentioned that the decision that had been asked about had been made by a decision maker other than the patient's personal general practitioner. Therefore, the general practitioner could not provide the requested information. Even though in Dutch health care general practitioners function as gatekeepers, this observation illustrates the fact that a single patient receives care from an increasing number of different health care workers. This increase in number of health care providers creates a need for a better management of health-care information.

In our study, two-thirds of the requests could be answered. To explain why the requested information had not been recorded (i.e., the information was available upon request), the general practitioners most frequently (54%) indicated that they normally did not record that information explicitly. For example, the motivation for a particular choice of therapy may not be recorded. Information about a physician's reasoning was often recorded implicitly, and available only when asked for[13, 20]. Some of the requested information turned out to be available elsewhere in the electronic medical record (17%). For example, information had been recorded as a short personal note in free text (not necessarily understood by others). In other cases, the information was assumed to be known by the readers of the medical record (22%). To address these limitations, the current electronic patient record will have to be modified with emphasis on structured data entry. The challenge that such systems have to face is to try to combine complexity with clarity and ease of use[21, 22].

The fact that requested additional information was available in many cases, supports the option to build an *inquisitive system*. About two-thirds of the missing information was available only when requested. However, an inquisitive system will have a much larger impact on the physician's normal routine, and therefore runs a larger risk of being rejected. Also, the majority of comments remained unchanged when the requested information became available, while we do not yet know the impact of the

minority of changed comments. Therefore, awaiting the results of our further studies, we have chosen a *non-inquisitive* design.

Having discussed the implications of our finding that critiquing comments could be made by human reviewers based upon data as they are currently available, we now discuss the kinds of comments that could be made and the general practitioners' assessments of these comments.

The largest categories of comments were those critiquing the prescribed medication and the general practitioner's therapeutic strategy in general. Interestingly, the reviewers' comments about the diagnostic phase of the patient-doctor encounter (though not made very frequently) were judged very positively. The general practitioners' positive response to these comments may suggest a need for support during the diagnostic phase. This observation seems to be in contradistinction to studies that have shown that diagnostic systems have had little impact on daily clinical practice[23]. Possibly, the kind of diagnostic support that is appreciated by physicians (support with diagnostic work-up) differs from the kind of support that diagnostic systems have provided in the past (support with differential diagnosis). When describing major obstacles to the implementation of decision-support systems, Taylor identified "loss of clinical control", as one of the possible reasons why diagnostic systems have achieved so little[9]. The fact that critiquing leaves the physician in control could account for our finding that general practitioners appreciated the diagnostic comments.

Prior to this study, we believed that if a general practitioner would disagree with the content of a critique, he would also judge that critique to be irrelevant. Overall, the agreement score and relevance score correlate with r=0.65. We were surprised to find that in a number of cases the general practitioner strongly disagreed with the content of a comment, but did not judge the comment to be irrelevant. This was most pronounced in the category of prescription-related comments. In other words, the general practitioners could see that a comment was relevant, but they could still strongly disagree with its content. This observation may imply that comments regarding prescriptions are very much needed from the point of view of the quality of health care – comments about prescriptions were frequently made– but that it will be a challenge to get physicians to accept prescription-related recommendations.

More insight is needed into the reasons why physicians reject critiquing comments in order to make the distinction between a reluctance of the physician to accept advice and a disagreement of the physician with the content of the advice.

CONCLUSION

We performed a simulation study of a computer system in order to gain insight into issues that determine the feasibility and effectiveness of an integrated critiquing system. Even though reviewers missed a considerable amount of information, our simulation study showed that it is possible for human reviewers (and therefore, theoretically feasible for computer algorithms) to generate critiquing comments based upon patient medical data as they are currently stored in electronic medical records. The largest categories of comments were about prescriptions and the physician's workplan. Comments regarding the diagnostic process are highly appreciated by the general practitioners. Interestingly, even though we investigated only a limited number of electronic medical records, the general practitioners judged prescription comments to be relevant, but often strongly disagreed with them. This discrepancy poses a challenge for the acceptability of critiquing comments that will be made by the future critiquing system. The general practitioners could provide answers to about two-third of the reviewers' requests for additional information. When this missing information was obtained, it led to changes in only a minority of generated comments. To provide integrated decision-support systems with more data than in our study, general practitioners' information systems will have to be developed that better support the structured entry of medical data. As a result of this study, we have started building the non-inquisitive critiquing system AsthmaCritic. AsthmaCritic will be subject to a field study, in which we will investigate the relationship between general practitioners' opinions of comments' correctness and relevance, the role of missing information, and the system's effectiveness.

ACKNOWLEDGMENTS

This study is supported by The Netherlands Asthma Foundation (#92.62).

The non-author reviewer in this study, Prof. E. van der Does, MD, PhD and the three general practitioners; J. Brienen, MD, C. Kunst, MD and J. van Wijngaarden, MD, receive special thanks.

REFERENCES

1. McDonald CJ. Protocol-based computer reminders, the quality of care and the non-perfectibility of man. N Engl J Med 1976;295:1351-5.
2. McDonald CJ, Wilson GA, McCabe GJJ. Physician response to computer reminders. JAMA 1980;244:1579-1581.
3. McDonald CJ, Hui SL, Smith DM, et al. Reminders to physicians from an introspective computer medical record. A two-year randomized trial. Ann Intern Med 1984;100:130-8.
4. Elson RB, Connelly DP. Computerized decision-support systems in primary care. Primary Care. 1995;22:365-84.
5. Shortliffe EH, Buchanan BG, Feigenbaum EA. Knowledge engineering for medical decision making: A review of computer-based clinical decision aids. Proc IEEE 1979;67:1207-24.
6. Heathfield HA, Wyatt J. Philosophies for the design and development of clinical decision-support systems. Meth Inform Med 1993;32:1-8.
7. Miller RA. Medical diagnostic decision support systems–past, present, and future: A threaded bibliography and brief commentary. J Am Med Informatics Assoc 1994;1:8-27.
8. Shortliffe EH. Testing reality: The introduction of decision-support technologies for physicians. Meth Inform Med 1989;28:1-5.
9. Taylor TR. The computer and clinical decision-support systems in primary care. J Fam Pract 1990;30:137-40.
10. Lei van der J, Duisterhout JS, Westerhof HP, et al. The introduction of computer-based patient records in The Netherlands. Ann Intern Med 1993;119:1036-1041.
11. Lei van der J, Musen M. A model for critiquing based on automated medical records. Comput Biomed Res 1991;24:344-78.
12. Lei van der J, Does van der E, Man in't Veld AJ, Musen MA, Bemmel van JH. Response of general practitioners to computer-generated critiques of hypertension therapy. Meth Inform Med. 1993;32:146-53.
13. Lei van der J, Musen M, Does van der E, Man in't Veld AJ, Bemmel van JH. Comparison of computer-aided and human review of general practitioners' management of hypertension. Lancet. 1991;338:1505-8.
14. Melker de RA, Jacobs HM, Kreuger FAF, Touw-Otten FWMM. Medische verslaglegging van huisartsen. Huisarts en Wetenschap 1994;37:46-51.
15. Gilliland AEW, Millis KA, Steele K. General practitioner records on computer–handle with care. Fam Pract 1992;9:441-50.
16. Shortliffe EHS, Perreault LE. Medical Informatics. Computer applications in health care. Reading, MA: Addison-Wesley, 1990.
17. Detmer WM, Shiffman S, Wyatt JC, Friedman CP, Lane CD, Fagan LM. A continuous-speech interface to a decision-support system: II. An evaluation using a Wizard-of-Oz experimental paradigm. J Am Med Informatics Assoc. 1995;2:46-57.
18. Overbeeke van JJ, Westerhof HP (eds). WCIA-HIS-Referentiemodel 1995.Volume A. Utrecht, The Netherlands: NHG/LHV Utrecht, NL, 1996:236.

19. Miller PL, Frawley SJ. Trade-offs in producing patient-specific recommendations from a computer-based clinical guideline: a case-study. J Am Med Informatics Assoc 1995;2:238-42.
20. Vlug AE, Lei van der J. Postmarketing surveillance with computer-based patient records. In: Greenes RA (ed). MEDINFO. Vancouver, BC, Canada: 1995: 327-30.
21. Ginneken van AM. Structured data entry in ORCA: the strengths of two models combined. In: Cimino JJ (ed). 1996 AMIA Annual Fall Symposium. Philadelphia, PA: Hanley&Belfus, 1996: 797-801.
22. Moorman PW, Ginneken van AM, Lei Van der J, Bemmel van JH. A model for structured data entry based on explicit descriptional knowledge. Meth Inform Med 1994;33:454-63.
23. Kassirer JP. A report card on computer-assisted diagnosis – the grade: C. N Engl J Med.

3

ASTHMACRITIC
ISSUES IN DESIGNING A NON-INQUISITIVE CRITIQUING SYSTEM FOR DAILY PRACTICE

Submitted for publication

Manon M. Kuilboer
Marc A. M. van Wijk
Mees Mosseveld
Johan van der Lei

ABSTRACT

To increase the acceptance of computer-based decision-support systems (CDSSs) in daily practice, the integration of such systems with the electronic patient record is highly advocated. We, therefore, chose to build a non-inquisitive critiquing system; a system that would use routinely recorded electronic patient data to select and analyze electronic patient records for the generation of critiquing comments. We designed the critiquing system reconciling the needs of a system functioning in physicians' busy daily routine. To implement the system we reused and expanded the generic critiquing system described by van der Lei. In this paper, we describe our design choices, we show how we reused the generic critiquing model to implement the system, we justify our design choices in light of existing literature and we summarize and reflect on issues underlying our design choices with respect to system acceptance.

INTRODUCTION

Asthma and chronic obstructive pulmonary disease (COPD) are chronic diseases with a high prevalence accounting for significant health-care expenditure[1]. Professional health-care organizations disseminate paper-based guidelines, which reflect the 'state of the art' in medical science[2,3]. Paper-based guidelines, however, have had disappointingly little impact on physicians' behavior[4,5]. Dissemination of guidelines alone is not enough to change daily practice; they need to be combined with an appropriate implementation strategy[6]. One such an implementation strategy is to introduce guidelines using computer-based decision-support systems (CDSSs)[7,8]. The objective of such systems is to help the practitioner manage patients with a particular disease using the appropriate guidelines and protocols. Although researchers have shown that computerized decision support is able to change health-care delivery, the number of systems in daily use is limited[7,9-14]. Some authors argue that the use of electronic patient records will provide new opportunities for decision support[11,15,16] - integration of decision-support facilities with the electronic patient record may provide a natural way to integrate decision support in every-day practice[7,14,17]. We designed and built a critiquing system in the domain of asthma and COPD, AsthmaCritic, taking integration into daily practice as a precondition. In this paper, we first describe AsthmaCritic's design and implementation. Next, we justify our choices in light of the available literature and denote issues underlying our choices. Finally, we summarize the issues that guided our design choices and discuss their implications.

ASTHMACRITIC, OVERVIEW

AsthmaCritic's task is to support the general practitioner with the diagnosis and treatment of patients with asthma or COPD during daily practice. In The Netherlands, most general practitioners use an electronic patient record to manage patient data[18]. The data are recorded by the general practitioner during consultation. Based on the data in the electronic record, AsthmaCritic provides feedback by generating a critique of the physicians' diagnostic and treatment plan. AsthmaCritic is a non-inquisitive critiquing system; the system does not ask for additional data entry. In order to deal with the constraints of a busy practice, the physician is always in control of AsthmaCritic's behavior. We will first give a brief overview of the system, followed by a description of the type of information provided by the system. Finally, we will discuss how the physician maintains control over the system's behavior.

AsthmaCritic has been integrated with a general practitioner information system. The system receives time-stamped patient data directly from the general practitioner's electronic patient record (Symptoms and diagnosis, Prescriptions, Measurements, Procedures, and Follow-up data). From the physician's viewpoint, AsthmaCritic is part of his/her medical record system. To emphasize the integration, the interface of AsthmaCritic is identical to the interface of the medical record system (that is, screen- and data-manipulation is handled in the same fashion). AsthmaCritic runs

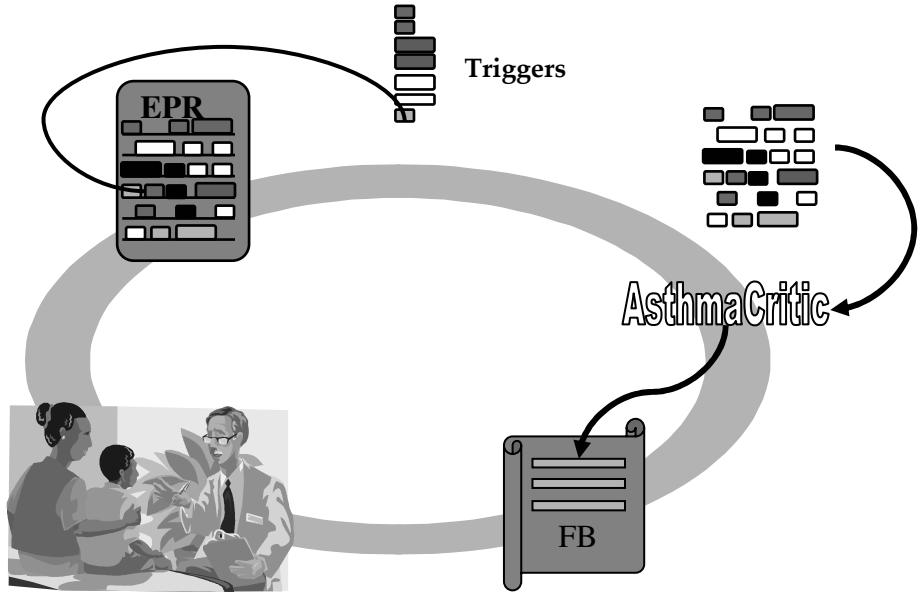


FIGURE 1

THE TRIGGERING OF ASTHMACRITIC. IF PATIENT DATA CORRESPOND WITH A SET OF PREDEFINED DATA CALLED TRIGGERS[1], THE SYSTEM WILL PERFORM AN ANALYSIS OF THE COMPLETE PATIENT RECORD. THIS ANALYSIS MAY LEAD TO THE GENERATION OF FEEDBACK. EPR = ELECTRONIC PATIENT RECORD. FB = FEEDBACK.

autonomously; the system is triggered when the physician sees an asthma or COPD patient[1], starts analyzing a record when the consultation is finished, and presents the critique (Figure 1 illustrates this process). The physician can interrupt the analysis of

1 Triggers used by AsthmaCritic: ICPC codes (' International Classification of Primary Care – coding system for Diagnosis, symptoms and procedures'[19]) for asthma (R96), chronic bronchitis (R91), emphysema (R95), other chronic pulmonary diseases (R83.4), and the ATC code for prescriptions (the Anatomical, Therapeutic and Chemical (ATC) coding system of the World Health Organisation[20]) used in the treatment of asthma or COPD (R03).

AsthmaCritic, thus forcing the system to return to the normal record-keeping mode. AsthmaCritic allows the physician to control *what* kind of feedback the system displays and *when* and *how* feedback has to be displayed. To provide this control to the physician, AsthmaCritic distinguishes three different kinds of feedback: Critiquing Information, Transformed Clinical Measurements, and a structured form of the Dutch guidelines (The Guideline Tree).

Critiquing information is presented to the physician as patient-specific comments based on the current clinical situation. AsthmaCritic first presents an overview of all comments. The comments are ranked depending on clinical urgency and the feedback's possible impact. For each comment, the physician can request additional information: a detailed description of the actual critique (including the source of the information), a description of the specific patient data that triggered the comment, and general information about the comment including references to literature.

Transformed clinical measurements constitute AsthmaCritic's second kind of information. Clinical measurements that are difficult to interpret are processed and presented using a layout tailored to their function. Peak flow values, for example, are presented to the physician in an overview that includes the original values, the expected value for this patient based on gender, age, and height, and the original values expressed as a percentage of the expected value. The physician may request such an overview upon his or her own initiative.

The Guideline Tree makes the guidelines of the Dutch Association of General Practitioners available in a structured and flexible form. For example, if a physician wants to check the current opinion about budesonide dosing schemas, he or she can find these schemas via the entry 'generic drug information'. However, a physician may have a patient having moderate asthma and may wonder whether budesonide is indicated. In such a case, the physician would search the desired information via the entry 'moderate asthma' instead of 'generic drug information'.

The general practitioner controls the behavior of AsthmaCritic. The physician decides when to review the feedback: during the patient encounter, at some time later, or at the patient's next visit. The physician also controls when the system runs its analysis: one hit on any key interrupts the analysis and initiates a background job that handles the record's processing. The physician may also decide to have AsthmaCritic always

analyze patient records in a background job during quiet moments of the day or night. If an analysis of a record is done in background mode, a patient-specific message is attached to the medical record that is displayed the next time the physician opens the record. The general practitioner's electronic patient record system keeps a log of the records that still have to be analyzed and a log of the feedback that still has to be read.

A description of the AsthmaCritic knowledge base is available to the physicians on paper. For quick reference on AsthmaCritic's functionality, a small laminated yellow reference card was available with on one side an overview of structured patient data relevant for asthma and COPD and on the other side the system's main control functions and helpdesk phone numbers. Finally, the system's manual also provided some background information on AsthmaCritic's functionality.

a. event descriptions	starting a drug
b. procedure	get all other active medication; get all diagnosis and symptoms; get all known interactions with the started drug; check for all those interactions whether the interacting drug is present and if so, whether the clinical effects of this interaction are true.
c. critiquing statement	interaction possible or present

FIGURE 2.

EXAMPLE OF A CRITIQUING TASK'S SPECIFICATIONS. IF A CLINICAL SITUATION AS DESCRIBED BY ITS EVENT DESCRIPTIONS, MATCHES A CRITIQUING TASK'S SPECIFICATIONS, FEEDBACK IS GENERATED.

ASTHMACRITIC, IMPLEMENTATION

AsthmaCritic's implementation is an extension of the generic critiquing model published by Van der Lei[21]. The generic critiquing model supports the integration with an electronic patient record at data level. The prototype, HyperCritic, however, was never tested in daily practice. HyperCritic lacked structures to enable the integration of the system in the physician's working environment. In addition, HyperCritic's ability to process time-stamped data was limited. Finally, HyperCritic presented its output as text, often several pages long, with no opportunity for the physician to control the behavior of the system. AsthmaCritic's implementation, therefore, differs from HyperCritic in increased use of time-stamped data, in distinguishing different types of information that allow the physician to control the output, and in supporting additional functions that allow integration in daily routine (such as, running the system in

background mode, attaching patient-specific messages to records, writing the results of analysis back into the medical record, or monitoring what feedback has been dealt with already).

CRITIQUING MODEL BASICS

The principle of the generic critiquing model is the generation of feedback based on a physician's treatment plan as reflected in patient data recorded in the electronic patient record. To generate feedback, the generic critiquing model makes a clear separation between *critiquing knowledge* (knowledge that initiates and guides the critiquing process) and *medical knowledge* (the medical base for critiquing). The analysis of patient data and medical knowledge is performed under control of the critiquing knowledge. The critiquing knowledge is described as a hierarchical set of critiquing tasks. Each individual critiquing task executes a specific procedure that defines for which clinical situations a critiquing statement should be generated. The bridge between patient data and the critiquing system is formed by events, e.g., 'starting budesonide' is an event that will be created if a physician records a new prescription of budesonide.

- Alternative drug task
- Consistent route of administration task
- Contraindication task
- Course verification task
- Dose/frequency task
- Incomplete prescription task
- Indication change therapy task
- Indication task
- Indication therapy task
- Interaction task
- Route of administration task
- Side-effect task
- Therapy trends task

FIGURE 3.

ASTHMACRITIC'S THIRTEEN DIFFERENT THERAPY-RELATED TASKS.

In order to trigger the critiquing tasks, *event descriptions* identify the criteria that have to be met by the patient data in order for the task to be executed. For example, the critiquing task that has as event description "starting a drug" will be triggered for any patient that starts any drug. Executing the critiquing task may result in a critiquing statement. Figure 2 shows a critiquing task that is executed whenever a drug is started. The task searches for interactions between drugs. If such an interaction is found, the critiquing statement "possible interaction" is generated[21].

The task specifications assume the existence of a knowledge base containing the relevant medical content. That is, critiquing tasks only specify the process of a critique, not the content. For the critiquing statement to be generated, medical knowledge needs to be available (for example, a drug hierarchy, dosing schedules, side-effects, interactions).

- Triggers (conditions)
  To define the clinical situation that needs to be true, e.g., age between 4 and 7 years=TRUE.
- Interpreted triggers (conditions)
  To define the patient's diagnostic state, e.g., diagnosis asthma = PROBABLE.
- Advice (conditions)
  To define the clinical situation that is adviced, e.g., prescribe inhaled corticosteroids =TRUE.
- Relevance(text)
  To define the comments rank in the Feedback Overview (three-point scale), e.g., 'very relevant'
- Short Title (text)
  To define a concise title used in Feedback Overview and the patient-specific message, e.g., 'Inconsistent route'.
- Title (text)
  To define the full title of critiquing statement, e.g., 'Inconsistent route of administration',
- Introduction to the advice (text)
  To define the text to start the critiquing statement with, e.g., 'The Dutch guidelines recommend to consider…'
- Data Source (text)
  To define the source of the knowledge, e.g., Dutch guidelines.
- Additional information (text with *hyperlinks*4).
  To define additional information for a comment e.g., an explanation of the relevance of a comment.

FIGURE 4.

A CRITIQUING TASK'S NINE DIFFERENT DATA STRUCTURES.

HyperCritic used only a limited set of event descriptions. In the more complex domain of asthma and COPD, we had to expand these event descriptions. For a complete

overview of AsthmaCritic's event descriptions, see Table 1. Unlike HyperCritic, AsthmaCritic creates event-histories from events. An *event-history* is a series of unique intervals of a class of event descriptions, for example, 'is-prescribing-drug X'. Within each interval, the patient's situation with respect to current prescriptions is stable. At the start or end of each interval, event descriptions specify the change of state. These event histories (one for drugs and one for measurements) enable the complex temporal analysis of the physicians' treatment plan in chronic medical diseases.

THE ASTHMACRITIC KNOWLEDGE BASES

For AsthmaCritic, we defined the following knowledge bases; the Critiquing knowledge base, the Medical knowledge base, the Supporting knowledge base, and the Educational knowledge base.

AsthmaCritic's knowledge base has been built from national guidelines[22-24], pharmaceutical reference books[25], guides on interactions and side effects[26], and consensus among a group of specialists in asthma and COPD. Building the knowledge base has been a 3-year iterative process under guidance of a medical content board consisting of four local specialists (two general practitioners and two pulmonologists[2] and seven national specialists[3]. Members of the medical content board reviewed each new version of the knowledge base.

CRITIQUING KNOWLEDGE

AsthmaCritic's Critiquing knowledge is divided into four categories of critiquing tasks: Diagnostic tasks, Therapy-related tasks, Referral-related tasks, and Follow-up-related tasks. Each category is further subdivided into more specific tasks. For example, *Therapy-related tasks* contains 13 different tasks, each tailored to a particular kind of clinical problem requiring specific data manipulation (Figure 3 shows the 13 tasks in alphabetical order).

In total, AsthmaCritic contains 131 specific tasks. These tasks, however, are solely procedural specifications. Based on the medical knowledge, the total number of different clinical situations that can be recognized is larger. For example, screening

2 B. Ponsioen, MD, E. van der Does, MD, PhD, J.C. de Jongste, MD, PhD, S.E. Overbeek, MD, PhD.
3 B. Bottema, MD, PhD, P.N.R. Dekhuyzen, MD, PhD, E.J. Duiverman, MD, PhD, E.E.M. van Essen, MD, PhD, Th. B. Voorn, MD, PhD, M.H.J. Vaessen, MD, A. van der Kuy, PhD.

for contra-indications is a single task, but the number of clinical situations that will be detected depend on the number of drugs and contra-indications in the medical knowledge.

For each specific critiquing task nine data structures can be used to store information. The data structures are used by the critiquing task mechanisms and by a Text Generator. The Text Generator generates critiquing statements using information from specific critiquing tasks. Figure 4 shows the nine data structures[4].

MEDICAL KNOWLEDGE

The Medical Knowledge Base provides the Critiquing Knowledge with basic medical concepts and their relationships. The concepts and their relationships are defined in a hierarchy of Concept classes and mechanisms, grouped by clinically meaningful topics. AsthmaCritic uses six different topics as shown in Figure 5.

- Medication knowledge,
  To define drug dosing schemas, route of administration, delivered units, KNMP code and ATC code,
- Measurement knowledge,
  To define normal values, ranges, and measurement unit,
- Problem knowledge,
  To define problems (coded (ICPC) and uncoded) and their characteristics, e.g., the default validity duration for a recorded problem,
- Relevance knowledge,
  To define each critiquing tasks' relevance on a three-point scale used by the interface, e.g., a comment 'Deteriorating peak flow' gets a relevance value 'extremely relevant',
- Specialist knowledge,
  To define clinical specialists patients may be referred to, e.g., the pulmonologist,
- Tag knowledge,
  To define tags used in the record, e.g.,'smoker'.

FIGURE 5.

ASTHMACRITIC'S SIX DIFFERENT TOPICS OF CONCEPT CLASSES USED IN THE MEDICAL KNOWLEDGE BASE

4 The *hyperlinks* are pointers to specific concept classes in the Guideline Tree.

EDUCATIONAL KNOWLEDGE

AsthmaCritic has available a body of medical information for educational purposes. Because of its different purpose, a separate hierarchical data structure was needed supporting flexible hierarchical access to its information. For AsthmaCritic, we structured the Dutch National Guidelines. We solved inconsistencies with our medical content board. In addition, the educational knowledge base is used by critiquing statements whose 'hyperlinks' may directly point to specific information in this knowledge base.

SUPPORTING KNOWLEDGE

In addition to offering (critiquing) information, AsthmaCritic implements functionality that supports the physician in interpreting complex clinical measurements. The system provides for a structure to store medical data needed for a transformation of such measurements and a specific mechanism to process the relevant events using domain-specific knowledge. For example, AsthmaCritic has implemented equations needed to calculate individual expected values for pulmonary measurements.

DESIGN CHOICES

We built a non-inquisitive critiquing system to support general practitioners with the diagnosis and treatment of patients with asthma and COPD in daily practice. Building the system, we made design choices aiming to optimize the system's chances to be accepted in the busy routine of general practice. In this section we will justify our main design choices using existing experience in literature. While doing so, we characterize the issues underlying those choices to be able to reflect on those issues in the discussion.

DATA-ENTRY, PHYSICAL LOCATION SINGULARITY, INTERFACE SINGULARITY, DATA-PROCESSING SPEEDNESS, AND APPLICATION CONSISTENCY

We chose to integrate AsthmaCritic with the general practitioner's information system. Integration of a computerized decision-support system with an electronic patient data source reduces the number of workflow interruptions[11,27-30]. Data entry itself should be a minimal nuisance - double, separate, complicated, and forced data entry are to be avoided (Issue: 'Data-Entry Effort')[31-33]. We, therefore, chose for a *non-inquisitive* system – a system that relies on routinely recorded data only and does not ask the user for additional data -. If integration with an electronic patient record is possible, a non-inquisitive system will reduce data-entry effort[35]. It is also known that acceptance

is limited if CDSSs don't provide their support on the same machine as the physician's information system (Issue: 'Physical location singularity'), don't share a common interface (Issue: 'Interface Singularity'), take a long time to process the data (Issue: 'Data-processing speedness'), and force the user to switch applications (Issue: 'Application Consistency')[34].

CASE-SUPPORT MATCHING

A related issue is the problem of matching patient records with proper guideline support (Issue: 'Case-Support Matching'). Realizing ourselves the additional hurdle physician-dependent matching would imply, we chose to let AsthmaCritic automatically select cases for further analysis. To select cases, AsthmaCritic uses specific data acting as *triggers*1. Other researchers also addressed this problem, e.g., by using Bayesian techniques or the Problem list[9,33]. The underlying idea is to automate the matching of cases and support, thereby avoiding physician-dependence.

PROFESSIONAL AUTONOMY, PROFESSIONAL INPUT, INFORMATION RELEVANCE

We chose for critiquing as the mode to offer support. A critiquing system provides a physician with feedback based on the physician's treatment plan as recorded in the patient's electronic patient record, after he or she made the decision. The first advantage of a critiquing system, therefore, is that it preserves the physician's professional autonomy by leaving him or her in control of the decision-making process (Issue: 'Professional Autonomy'). Already in 1989, Shortliffe described 'loss of control' as one of the physician-perceived barriers to the introduction of decision-support systems. This sense of loss of control was perhaps in part inspired by the fear physicians had about 'expert systems' taking over their jobs, which lead to a rejection of the decision-support concept altogether[36,37]. Taylor, in 1990, described another 'loss of control', which was about the psychology of being denied the reward of using one's own skills (Issue: 'Professional Input')[38]. The second advantage of the critiquing system approach, is that the risk for physicians to become dependant on CDSSs is smaller, precisely because physicians first have to make their own decision. If physicians become dependent on a CDSSs, decisions may run a larger risk being sub-optimal because, for example, physicians become less alert to abnormal conditions. The third advantage is that critiquing comments are patient-specific, having a higher chance of being relevant for the situation at hand. Physicians are extremely critical about the relevance of available information because time is short (Issue: 'Information Relevance').

INFORMATION DIFFERENTIATION

We chose to make explicit the different kinds of information generated in feedback. This choice deals with several issues; the problem of limited time, a (relative) information overload, and the variability of available time and information needs[39,40]. In daily practice, there is no time for users to read all available information in order to identify different kinds of information and select what they need[41]. Choosing is possible if differences in information have been made explicit (Issue: 'Information Differentiation').

INFORMATION CONCISENESS

We took care to limit the amount of information presented at each instance. If physicians are expected to read and process information when time is limited, the amount of information has to be limited (Issue: 'Information Conciseness')[42]. Lobach found that the physicians in his study preferred a clear telegraphic style when it comes to the presentation of guideline info - large text bodies were unwanted.

INFORMATION JUSTIFICATION

We chose to present with feedback, information about why the feedback had been generated and where the information had been based on (Issue: 'Information Justification'). Previous experience has shown that users' trust in generated advice is increased if the system is able to justify its recommendations[43,44]. Neural networks constitute an example of systems that partly failed, not only because they ignore physicians' professional and personal autonomy, but also because they function as a 'black box' with no options for the user to follow the neural network's reasoning. To further increase physicians' trust in AsthmaCritic, we chose to provide physicians also with paper-based information describing the system's knowledge base in a user-friendly manner.

TIMING AND PERSONAL AUTONOMY

We chose to give the physician complete control over timing and content of information exposure (Issue: 'Timing'). Timing and freedom are important issues; physicians should be able to read or interrupt reading generated support at any moment (Issue: 'Personal Autonomy')[45]. Experience with PROMIS already showed that forcing physicians into a specific structure might lead to rejection[31]. Providing the physician with control is one way to match feedback with physicians' time and info needs.

DESIGN CHOICES DISCUSSED

The issues underlying our design choices deal with hurdles the physician has to overcome in two different user phases. First, issues that deal with the hurdles the user has to overcome to enable the critiquing system to generate the output – the user phase *Generating Output*. For example, a physician having to record patient data specifically for a support program ('Data-entry effort') or a user having to select by himself the support that is proper for the situation at hand ('Case-support matching'). Second, issues that deal with the hurdles the physician has to overcome to be able to use the feedback – the user phase *Using Output*. For example, whether an amount of text presented is quick and easy to understand ('Information Conciseness'), or whether the timing of feedback fits the user's moment of interest ('Timing'). Table 2 summarizes the issues and our design choices regarding each of them.

The difference between the issues in the two user phases is, that with issues playing a role during 'Generating output', user involvement should be *minimized* in order to optimize user acceptance. The physician should, for example, not be bothered with providing the system with patient data and getting it started to run an analysis. With issues playing a role during 'Using Output', it is the other way around. User involvement should be *maximized* in order to optimize user acceptance. Physicians, for example, have to be able to interrupt an analysis if they have no time, and should be able to get further detail on why comments had been generated if they feel the need to know. During the phase of 'Generating output', therefore, control by the physician is unwanted, while during 'Using output', control by the physician is required.

As system developers we had to make design choices for each of the described issues. To make our own choices, we had to characterize general practitioners' working environment, which we could only do in broad terms. What we missed was an established understanding of the relationship between the characteristics of CDSSs and the characteristics of different working environments with respect to system acceptance. For each issue, a system developer should know how each choice influences system operation. If he or she knows these relationships, design choices can be optimized considering the constraints and options of the system's intended working environment. We feel that further insight into the relationship

between CDSSs and working environments is needed. This insight will be helpful for system designers and researchers alike, and will hopefully reduce design errors.

For one of our own design choices (regarding the issue 'Data-Entry Effort') we chose to build a non-inquisitive system. AsthmaCritic would be integrated with the general practitioner information system and receive its patient data straight from the information system. It would not interrupt the user to request specific or additional data. However, the availability of structured medical data depends on physicians' ability and willingness to record data in a structured fashion. These recording habits are highly variable. The question is whether, given this variability, the choice for non-inquisitiveness is feasible. Therefore, in a previous study, we investigated the feasibility of using data routinely recorded in electronic patient records for the generation of patient-specific feedback[35]. We concluded that enough structured data were available to generate relevant feedback for general practitioners. In the study, we took one step further by investigating the need to ask physicians for missing data[47]. Our study revealed that information that was being missed by reviewers was, very often, available elsewhere, but did not make much difference in the generation of the comments. Therefore, aiming to minimize Data-Entry Effort, and given the availability of a general practitioner information system, we decided building a non-inquisitive critiquing system.

Finally, for the implementation of AsthmaCritic we reused the generic critiquing model published by Van der Lei[21]. The generic critiquing model supports the integration with an electronic patient record at data level. The prototype, HyperCritic, however, was never tested in daily practice and lacked structures to enable the functional integration of the system in the physician's working environment. AsthmaCritic's implementation, therefore, differs from HyperCritic in supporting additional functions that allow such an integration in daily routine. In other words, the generic critiquing model had to be expanded to accommodate the requirements of a system for daily practice. However, the model partially fit our needs and thus proved to be reusable.

ACKNOWLEDGMENTS

We thank our local medical content board: B. Ponsioen, MD, prof. E. van der Does, MD, PhD, Prof. J. C. de Jongste, MD, PhD, S. E. Overbeek, MD, PhD and our national medical content board: B. Bottema, MD, PhD, P. N. R. Dekhuyzen, MD, PhD, E. J. Duiverman, MD, PhD, E. E. M. van Essen, MD, PhD, Th. B. Voorn, MD, PhD, M. H. J. Vaessen, MD, and A. van der Kuy, PhD for their part in the development of AsthmaCritic knowledge base.

We also thank M. Kroneman, MSc, PhD for her valuable contributions to this paper.

This study is supported by The Netherlands Asthma Foundation (#92.62).

TABLE 1.

ASTHMACRITIC EVENT DESCRIPTIONS.

EVENT DESCRIPTION
Is-stopping drug
Is-starting drug
Is-decreasing drug
Is-increasing drug
Is-prescribing drug
Prescribed drug in the past
Prescription frequency-is-smaller than
Prescription frequency-is-larger than
Prescription frequency equals
Measurement is-increasing
Measurement is decreasing
Measurement is instable
Measurement is normal
Measurement value equals value
Measurement value equals value in the past
Measurement value is smaller than value
Measurement value is larger than value
Measurement fraction equals fraction
Measurement fraction equals target fraction
Measurement fraction is smaller than fraction
Measurement fraction is smaller than target fraction
Measurement fraction is smaller than target fraction in the past
Measurement fraction is larger than fraction
Measurement fraction is larger than target fraction
Measurement fraction is larger than target fraction in the past
Trend two latest measurements is decreasing
Trend two latest measurements is increasing
Trend two latest measurements is stable
Patient age equals
Patient age is younger than
Patient age is older than
Patient age is between
Is having a symptom/diagnosis (frequency, period)
Is being referred
Is having a tag
Test has been performed
Is happening in time of the year

TABLE 2.

ISSUES PLAYING A ROLE DURING TWO USERS' PHASES; 'GENERATING OUTPUT' AND 'USING OUTPUT', INCLUDING OUR DESIGN CHOICES REGARDING EACH OF THESE ISSUES.

	ISSUE	ASTHMACRITIC
Generating Output	Application singularity	Is integrated with the general practitioner information system (ELIAS)
	Case-support matching	Automates case-support matching by the use of triggers
	Data-entry effort	Is non-inquisitive – it receives routinely recorded patient data from the information system and does not ask for specific data entry
	Data-processing speedness	Bypasses the MEDEUR[48] interface.
	Physical location singularity	Runs on the same machine as the general practitioner information system (ELIAS)
Using Output	Information conciseness	Presents information in layered text bodies, following a strict hierarchy
	Information differentiation	Makes differences in information explicit
	Information justification	Shows patient data leading to the generation of feedback upon request
	Information relevance	Critiques the physician on a patient-specific level; Provides the user with the tools to control information exposure
	Interface consistency	Matches the general practitioner information system (ELIAS) interface conventions
	Personal autonomy	Lets the user start or interrupt processing or reading at any moment
	Professional autonomy	Allows the user to make the medical decisions
	Professional input	Takes the users' actions as the base for critiquing, and requires the user to apply his own knowledge in the interpretation of offered feedback
	Timing	Starts analyzing right after the last patient data have been recorded

REFERENCES

1 Rutten-van Molken, M.P. et al. (1999) Current and future medical costs of asthma and chronic obstructive pulmonary disease in The Netherlands. *Respir Med* 93 (11), 779-787

2 Podell, R.N. (1992) National guidelines for the management of asthma in adults. *Am Fam Physician* 46 (4), 1189-1196.

3 Geijer, R.M.M. et al. (1997) NHG-Standaard COPD en astma bij volwassenen: Diagnostiek [Guidelines of the Dutch College of General Practitioners: Chronic Obstructive Respiratory Diseases and Asthma in Adults: Diagnostics]. *Huisarts en Wetenschap* 40 (9), 415-428

4 Crim, C. (2000) Clinical practice guidelines vs actual clinical practice. The asthma paradigm. *Chest* 118, 62S-64S

5 Smeele, I.J. et al. (1998) [Discrepancy between the guidelines and practice by family physicians in treating adults with an exacerbation of asthma or chronic obstructive pulmonary disease]. *Ned Tijdschr Geneeskd* 142 (42), 2304-2308.

6 Forrest, D. et al. (1996) Clinical guidelines and their implementation. *Postgrad Med J* 72 (843), 19-22

7 Hunt, D.L., MD et al. (1998) Effects of computer-based clinical decision support systems on physician performance and patient outcomes: A systematic review. *Journal of the American Medical Association* 280, 1339-1346

8 van Wijk, M. et al. (1999) Design of a decision support system for test ordering in general practice: choices and decisions to make. *Methods Inf Med* 38 (4-5), 355-361.

9 Lobach, D.F. and Hammond, W.E. (1997) Computerized decision support based on a clinical practice guideline improves compliance with care standards. *Am J Med* 102 (1), 89-98

10 Haynes, R.B. et al. (1996) Transferring evidence from research into practice: 1. The role of clinical care research evidence in clinical decisions [editorial]. *ACP J Club* 125 (3), A14-16

11 Elson, R.B., MD and Connelly, D.P., MD, PhD. (1995) Computerized decision-support systems in primary care. *Primary Care* 22 (2), 365-384

12 Mitchell, E. and Sullivan, F. (2001) A descriptive feast but an evaluative famine: systematic review of published articles on primary care computing during 1980-97. *British Medical Journal* 322 (7281), 279-282.

13 Zielstorff, R.D. (1998) Online practice guidelines: issues, obstacles, and future prospects. *J Am Med Inform Assoc* 5 (3), 227-236

14 Johnston, M.E., BSc et al. (1994) Effects of computer-based clinical decision-support systems on clinician performance and patient outcome: A critical appraisal of research. *Annals of Internal Medicine* 120, 135-142

15 Tierney, W.M. et al. (1995) Computerizing guidelines to improve care and patient outcomes: the example of heart failure. *J Am Med Inform Assoc* 2 (5), 316-322

16 Henry, S.B. et al. (1998) A template-based approach to support utilization of clinical practice guidelines within an electronic health record. *J Am Med Inform Assoc* 5 (3), 237-244

17 Smith, B.J. and McNeely, M.D. (1999) The influence of an expert system for test ordering and interpretation on laboratory investigations. *Clin Chem* 45 (8 Pt 1), 1168-1175

18 Lei van der, J., MD, PhD et al. (1993) The introduction of computer-based patient records in The Netherlands. *Annals of Internal Medicine* 119 (10), 1036-1041

19 Boersma, J.J. (1995) ICPC: International classification of primary care: short titels and Dutch subtitels, Nederlands Huisartsen Genootschap

20 WHO International working group for drug statistics methodology. (1996) Anatomical therapeutic chemical (ATC) classification and defined daily dose (DDD) for pharmaceuticals and vitamins. At: www.whocc.nmd.no/ Last accessed:

21 Lei van der, J. and Musen, M.A. (1991) A model for critiquing based on automated medical records. *Computers and Biomedical Research* 24, 344-378

22 Dirksen, W.J. et al. (1992) NHG Standaard Astma bij Kinderen [Guidelines of the Dutch College of General Practitioners: Asthma in Children]. *Huisarts en Wetenschap* 35 (9), 355-362

23 Bottema, B.J.A.M. et al. (1992) NHG Standaard CARA bij Volwassenen: Diagnostiek [Guidelines of the Dutch College of General Practitioners: Chronic Respiratory Diseases in Adults: Diagnostics]. *Huisarts en Wetenschap* 35 (11), 430-436

24 Waart van der, M.A.C. et al. (1992) NHG Standaard CARA bij Volwassenen: Behandeling. [Guidelines of the Dutch College of General Practitioners: Chronic Respiratory Diseases in Adults: Therapy]. *Huisarts en Wetenschap* 35 (11), 437-443

25 Kuy van der, A., ed. (1997) Farmacotherapeutisch Kompas [Pharmaceutical Reference Book], Boekhoven-Bosch, Utrecht

26 Akkerveen van, H.M., Drs et al. (1994) Commentaren Medicatiebewaking Pharmacom en Medicom [Comments in Medication Alerts Pharmacom and Medicom], Stichting health base

27 Lobach, D.F., MD PhD and Underwood, H.R., MD MBA. (1998) Computer-based decision support systems for implementing clinical practice guidelines. *Drug Benefit Trends* 10 (10), 48-53

28 Leader, W.G. et al. (1996) Integrating pharmacokinetics into point-of-care information systems. *Clin Pharmacokinet* 31 (3), 165-173

29 Gadd, C.S. et al. (1998) Identification of design features to enhance utilization and acceptance of systems for Internet-based decision support at the point of care. *Proc AMIA Symp*, 91-95

30 Shiffman, R.N., MD, MCIS. (1994) Towards effective implementation of a pediatric asthma guideline: integration of decision support and clinical workflow support. In *Proceeding of the Annual Symposium on Computer Application in Medical Care* (Vol. 1994) (Ozbolt, J.G., PhD, RN, ed.), pp. 797-801, Hanley & Belfus, Inc.

31 Wiederholt, G. and Perreault, L.E. (1990) Hospital Information Systems. In *Medical Informatics. Computer applications in health care* (Shortliffe, E.H.S. and Perreault, L.E., eds.), pp. 237, Addison-Wesley Publishing Company, Inc.

32 Porcelli, P.J. and Lobach, D.F. (1999) Integration of clinical decision support with on-line encounter documentation for well child care at the point of care [In Process Citation]. *Proc AMIA Symp*, 599-603

33 Aronsky, D. and Haug, P.J. (1999) An integrated decision support system for diagnosing and managing patients with community-acquired pneumonia [In Process Citation]. *Proc AMIA Symp*, 197-201

34 Ebell, M.H. et al. (1997) Family physicians' preferences for computerized decision-support hardware and software [see comments]. *J Fam Pract* 45 (2), 137-141

35 Kuilboer, M.M. et al. (1998) Exploring the role of an integrated critiquing system: A simulation. *Journal of the American Medical Informatics Association* (5), 194-202

36 Shortliffe, E.H. (1989) Testing reality: The introduction of decision-support technologies for physicians. *Methods of Information in Medicine* 28 (1), 1-5

37 Miller, R.A. and Masarie, F.E., Jr. (1990) The demise of the "Greek Oracle" model for medical diagnostic systems. *Methods of Information in Medicine* (29), 1-2

38 Taylor, T.R., MD, PhD. (1990) The computer and clinical decision-support systems in primary care. *Journal of Family Practice* 30 (2), 137-140

39 Tang, P.C. et al. (1995) Methods for assessing information needs of clinicians in ambulatory care. *Proc Annu Symp Comput Appl Med Care*, 630-634

40 Fafchamps, D. et al. (1991) Modelling work practices: input to the design of a physician's workstation. *Proc Annu Symp Comput Appl Med Care*, 788-792

41 Noone, J. et al. (1998) Information overload: opportunities and challenges for the general practitioner's desktop. *Medinfo* 9 (Pt 2), 1287-1291

42 McDonald, C.J. (1976) Protocol-based computer reminders, the quality of care and the non-perfectibility of man. *New England Journal of Medicine* 295 (24), 1351-1355

43 Teach, R.L. and Shortliffe, E.H. (1981) An analysis of physician attitudes regarding computer-based clinical consultation systems. *Comput Biomed Res* 14 (6), 542-558.

44 Morris, A.H. (2000) Developing and implementing computerized protocols for standardization of clinical decisions. *Ann Intern Med* 132 (5), 373-383.

45 Tierney, W.M. et al. (1986) Delayed feedback of physician performance versus immediate reminders to perform preventive care. Effects on physician compliance. *Med Care* 24 (8), 659-666

46 Grol, R. et al. (1995) Development and implementation of guidelines for family practice: lessons from The Netherlands [editorial]. *J Fam Pract* 40 (5), 435-439

47 Kuilboer, M.M. et al. (1997) The availability of unavailable information. In *1997 AMIA Annual Fall Symposium*, pp. 749-753, Hanley&Belfus, Inc.

48 Vlug, A. MEDEUR homepage. At: http://www.eur.nl/fgg/mi/medeur/ Last accessed: 14 07 2002.

4

FEASIBILITY OF ASTHMACRITIC, A DECISION-SUPPORT SYSTEM FOR ASTHMA AND COPD, WHICH GENERATES PATIENT-SPECIFIC FEEDBACK ON ROUTINELY RECORDED DATA IN GENERAL PRACTICE

Published in Family Practice; 2002; 19 (5): 442-447

Manon M. Kuilboer
Marc A. M. van Wijk
Mees Mosseveld
Emiel van der Does
Ben P. Ponsioen
Johan C. de Jongste
Shelley E. Overbeek
Johan van der Lei

ABSTRACT

BACKGROUND

Introducing decision-support systems as a tool to stimulate the dissemination of clinical guidelines in daily practice has been disappointing. Researchers have argued that integration of such systems with clinical practice is a prerequisite for acceptance. The big question concerns the feasibility of a true integration – if only routinely recorded data are used for such a system, can patient-specific feedback be produced?

OBJECTIVE

To assess the feasibility of generating patient-specific feedback based on routinely recorded data in general practice by AsthmaCritic, a decision-support system for asthma and chronic obstructive pulmonary disease (COPD).

METHODS

We built the decision-support system AsthmaCritic. We assessed AsthmaCritic's ability to detect asthma and COPD patient records and generate patient-specific feedback. We grouped feedback into categories of comments by age group (<12 years and ≥ 12 years).

DESIGN

Retrospective analysis of routinely recorded data in 103,713 electronic patient records from primary-care practices.

MAIN OUTCOME MEASURES

Number and percentage of 'triggered' (selected) asthma and COPD patient records. Number and percentage of records on which AsthmaCritic produced at least one feedback comment during the one-year study period, by category of comments.

RESULTS

AsthmaCritic detected 8784 (8.5%) asthma and COPD patient records. During the study period, AsthmaCritic generated 255,664 feedback comments (mean 3.4 per patient visit). The most frequently generated category of comments in case of patients12 years or older, was *Non-compliant Prescription* (23.7%) whereas the most frequent category in case of patients younger than 12 years was *Non-compliant Route* (31.1%).

CONCLUSIONS

This study shows that, using routinely recorded data only, AsthmaCritic is able to detect asthma and COPD patient records for further analysis and to produce patient-specific feedback.

INTRODUCTION

Asthma and COPD are chronic diseases with a high prevalence accounting for significant health-care expenditure[1]. In recent years, the treatment of asthma and COPD has changed considerably. The consecutive guidelines for asthma and COPD issued by the Dutch College of General Practitioners, for example, illustrate the development of new treatment regimens[2-8]. Physicians face the challenge of coping with the changing and ever-increasing amount of medical knowledge[9-11]. In view of the current emphasis on evidence-based medicine, clinical practice guidelines[12] are considered to be an important tool for disseminating new medical knowledge[13-15]. Nevertheless, their use in daily practice has been disappointingly low[16-20]. Computer-based decision-support systems may facilitate the implementation of guidelines in daily practice[21, 22]. However, to be successful, many investigators argue that these systems need to be integrated with computer-based patient records[23-26]. In the absence of such integration, physicians have to record data already available in the electronic medical record a second time.

In The Netherlands, most general practitioners have replaced their paper-based patient records with computer-based records; the practitioners themselves record patient data into the computer during patient encounters[25]. To code patient data, they use the International Classification of Primary Care (ICPC) for symptoms, procedures, and diagnosis[27]. Prescriptions are coded according to the Anatomical, Therapeutic and Chemical (ATC) coding system of the World Health Organisation[28]. The general practitioner may also record data as free text.

As the first, essential step to demonstrate the feasibility of integrated support, we developed AsthmaCritic, a computer-based decision-support system for asthma and COPD, and let it analyse routinely recorded data in electronic patient records of general practitioners. In this paper, we first describe the system, followed by a description and discussion of our feasibility study.

ASTHMACRITIC

The objective of the decision-support system AsthmaCritic is to review the physician's treatment in the light of the most recently published guidelines. The system generates patient-specific feedback in the form of critiquing comments. These comments review

the physician's diagnostic and therapeutic interventions thus enabling physicians to reflect on their decisions, while being focussed on the patient at hand. AsthmaCritic generates these comments based on data routinely recorded by the general practitioner in an electronic patient record.

The knowledge base of AsthmaCritic is predominantly derived from the asthma and COPD guidelines of the Dutch College of General Practitioners[2-4]. Building the knowledge base has been a 3-year iterative process under guidance of a *medical content board* consisting of four local experts (two general practitioners, BP and ED, a pulmonologist, SO, and a paediatric pulmonologist, JJ) and seven national experts. Members of the medical content board reviewed each new version of the knowledge base.

At the end of each patient contact, the electronic patient record activates AsthmaCritic. AsthmaCritic first searches the medical record for clues, *triggers,* that indicate the possibility of asthma or COPD: ICPC codes for asthma (R96), chronic bronchitis (R91), emphysema (R95), other chronic pulmonary diseases (R83.4), and the ATC code for prescriptions used in the treatment of asthma or COPD (R03). When AsthmaCritic encounters a trigger, the record is selected for a full analysis. AsthmaCritic subsequently reviews different aspects of the physician's treatment and may generate feedback. The system does not question the correctness of the data recorded by the physician. For example, if the physician records a diagnosis asthma, AsthmaCritic does not judge the physician's opinion.

AsthmaCritic presents feedback to the general practitioner as a list of brief comments. The system is able to provide for each comment one or more of the following kinds of additional information: an elaborated advice, a further explanation, the applied patient data, or the underlying medical knowledge. By selecting a comment ("clicking the comment"), the general practitioner can access the additional information. If, for example, the system detects a decrease in peak flow, a short comment "*Alarming situation: decreasing peak flow*" is included in the list; by selecting that comment, the general practitioner can inspect the elaborated advice, the patient data, the interpretation of the measurement, and the relevant sections of the guidelines.

METHODS

STUDY DESIGN

To assess the feasibility of our approach, we analysed electronic patient records of over 100,000 patients in 28 general practices. This analysis consisted of two stages. AsthmaCritic first examined all complete records to detect triggers, that is, the identification of data pointing to asthma or COPD; records containing a trigger, the so-called triggered records, were marked for further analysis. Of the triggered records, the system subsequently reviewed each patient contact within the study period (January 1996 through December 1996). Reflecting different aspects of treatment, we divided AsthmaCritic's comments into twelve *categories*; Table 1 shows a short description and a brief example for each category. The category *alarming situations*, for example, are those comments that detect a deterioration of the patient's condition. Adhering to the Guidelines of the Dutch College of General Practitioners, we divided the population into two age groups; one including patients younger than twelve years, and one including patients twelve years and older[2, 3].

SETTING

The Department of Medical Informatics of the Erasmus Medical Center Rotterdam collaborates with general practitioner practices located in different parts of the country that make their data available for research in primary care[29]; in 1996 – the study period –, the number of collaborating practices was 28. From these practices we retrieved the electronic patient records of all patients enrolled in these practices in 1996; these records were subsequently analysed by AsthmaCritic.

MEASUREMENTS

We counted the number and calculated the percentage of AsthmaCritic's triggered records. For the triggered records, we counted the number of comments, the number of contacts, and we calculated the average number of comments per contact. For the different categories of comments, we calculated the percentage of triggered records in which at least one comment from that category was made during the study period (counting each instance of a generated comment would yield unrealistic frequencies because of the retrospective nature of the study – physicians could not change their behaviour in response to generated comments, therefore, once a comment was generated and the circumstances did not change, a comment was generated at each contact).

TABLE 1.

CHARACTERIZATION OF CATEGORIES OF COMMENTS.

CATEGORY	DESCRIPTION	EXAMPLE
Alarming situations	Signs of deterioration	A decrease in peakflow
Change in therapy advised	Changes in medication recommended	Start a short course oral corticosteroids
Contraindications	Contraindication present	Known NSAID sensitivity
Dose deviations	Non-compliant dose	Dose lower than recommended
Frequency deviations	The dose frequency is non-compliant	More doses per day than recommended
Non-compliant route	The route of administration deviates from the guidelines	A powder inhaler in a three-year old child
Inconsistent route	Multiple different inhaler devices prescribed	A metered dose inhaler combined with a powder inhaler
Non-compliant prescriptions	Medication is prescribed as “on demand” or “fixed” in contrast with the guidelines	Inhaled corticosteroids are recommended to be prescribed “on demand”
Interactions	Possible interactions between different drugs	Chinolones and xanthine derivatives may interact and decrease metabolic clearance, causing nausea, vomiting, headache and/or vertigo
Early Reduction	Therapy is reduced sooner than recommended	Reduction of inhaled steroid within 2 weeks
Side effects	Side effect detected	Thrush with inhaled corticosteroids
Many antibiotics	Frequent courses of antibiotics	Frequent prescription of antibiotics without having started a course with corticosteroids

RESULTS

During the study period, 103,713 patients were enrolled in the 28 practices. Of the 103,713 records, 8784 (8.5%) were selected by AsthmaCritic for further analysis: 53.6% were triggered by diagnosis and 46.4% by medication. Of the 8784 patients with a trigger in their record, 8412 had at least one encounter with the general practitioner during the study period. Of the 8412 patients with at least one encounter, 6190 (73.6 %) were 12 years and older (3352 female), and 2222 (26.4 %) were younger than 12 years (1005 girls). An overview of the results is presented in Table 2.

TABLE 2.

DESCRIPTIVE STATISTICS OF TRIGGERED RECORDS BY AGE GROUP (TOTAL PATIENT POPULATION; N=103,713).

	≥12 YEARS	<12 YEARS	TOTAL
Number of triggered records			8784
Number of triggered records with >= 1 contact	6190 (73.6%)	2222 (26.4%)	8412
Number of males	2838 (45.8%)	1217 (54.8%)	4055
Number of females	3352 (54.2%)	1005 (45.2%)	4357
Number of comments	237179	18485	255664
Number of contacts	62389	12320	74709
Average number of contacts/triggered record	10	5.5	9
Average number of comments/contact	3.6	1.5	3.4

Of the 8412 patients who had at least one encounter with their general practitioner in 1996, AsthmaCritic performed an analysis of all encounters during 1996, taking all information preceding each encounter into account. The 8412 patients who had seen their general practitioner in 1996 had a total of 74,709 encounters with their general practitioner; an average of 9 contacts per patient (patients aged 12 years and older had an average of 10 contacts, mode: 5, SD: 9; and patients aged younger than 12 years had an average of 5.5 contacts, mode: 2, SD: 4). AsthmaCritic reviewed all 74,709 encounters in 1996 and generated in total 255,664 comments, an average of 3.4 comments per encounter. For the different categories of comments, we calculated

the percentage of the triggered records in which at least one comment from that category was made during the study period. The results for patients aged 12 years and older are shown in Table 3, for patients aged younger than 12 years in Table 4. The most frequently generated category of comments in patients aged 12 years and older was *Non-compliant Prescription* (of the 6190 triggered records, 1467 (23.7%) at least once in the study period) whereas the most frequent category in patients aged younger than 12 years was *Non-compliant Route* of administration (31.1%).

TABLE 3.

FOR PATIENTS ≥ 12 YEARS (N=6190), THE NUMBER AND PERCENTAGE OF RECORDS IN WHICH AT LEAST ONE OF THE COMMENTS OF A GROUP OF COMMENTS HAD BEEN GENERATED.

PATIENTS ≥ 12 YEARS **GROUP OF COMMENTS**	**NUMBER OF RECORDS**	**PERCENTAGE OF RECORDS**
Non-compliant prescriptions	1467	23,7%
Contraindications	1381	22,3%
Alarming situations	912	14,7%
Dose deviations	683	11,0%
Inconsistent route	598	9,7%
Many antibiotics	534	8,6%
Early Reduction	444	7,2%
Change in therapy advised	381	6,2%
Frequency deviations	350	5,7%
Interactions	175	2,8%
Non-compliant route	101	1,6%
Side effects	43	0,7%

TABLE 4.

FOR PATIENTS < 12 YEARS (N=2222), THE NUMBER AND PERCENTAGE OF RECORDS IN WHICH AT LEAST ONE OF THE COMMENTS OF A GROUP OF COMMENTS HAD BEEN GENERATED.

PATIENTS < 12 YEARS GROUP OF COMMENTS	NUMBER OF RECORDS	PERCENTAGE OF RECORDS
Non-compliant route	691	31,1%
Non-compliant prescriptions	304	13,7%
Dose deviations	288	13,0%
Change in therapy advised	273	12,3%
Alarming situations	236	10,6%
Frequency deviations	167	7,5%
Many antibiotics	136	6,1%
Inconsistent administration	79	3,6%
Early Reduction	67	3,0%
Contraindications	20	0,9%
Interactions	4	0,2%
Side effects	4	0,2%

DISCUSSION

Integrating decision-support systems with electronic patient records is an important factor in the applicability of such systems in daily practice[23, 24]. In a previous study, we showed that electronic patient records contain sufficient information for experts to review the treatment of asthma and COPD [30]. Based on this study, we built AsthmaCritic, a system that generates critiquing comments using data routinely recorded by general practitioners in their electronic patient records. In this study, AsthmaCritic selected 8.5% of over 100,000 records as belonging to patients with asthma or COPD, which matches with the 5 to 10 % prevalence rate known from Dutch registration networks[5-8, 31-33]. Of the selected records, AsthmaCritic analysed the medical record for each of the 74.709 encounters, and generated a total of 255.664 comments, an average of 3.4 per encounter.

For patients aged older than 12 years, the most frequent comment of AsthmaCritic was the category *Non-compliant Prescriptions* (23.7%). Although comments in the category *Non-compliant Prescriptions* are also frequent in patients aged younger than 12 years (13.7 %), the most frequent comment in patients aged younger than 12 years was the category *Non-compliant Route* (31.1%). Compared to the guidelines for patients aged 12 years and older, determining the optimal route of administration is difficult in patients aged younger than 12 years; the route depends on age and the patient's clinical condition[34]. It is, therefore, not surprising that comments in the category *Non-compliant Route* are much more frequent in patients aged younger than 12 years than in patients aged 12 years and older.

Because decision-support systems regard data with a limited scope, physician interpretation of comments will be needed to determine AsthmaCritic's clinical relevance. The extent to which physician judgement is required depends on a comment's category. For example, in 22.3% of the patients aged older than 12 years, AsthmaCritic pointed out the presence of contraindications. Many of these contraindications, however, are relative. AsthmaCritic will point out that asthma is a contraindication for the prescription of cyclo-oxygenase inhibitors. The physician, however, may accept that risk. Another example that underscores the importance of physician interpretation is comments dealing with the frequent use of antibiotics. AsthmaCritic will generate comments when the patient receives four or more courses of antibiotics over a period of twelve months. In 8.6 % of the patients aged 12 years and older and 6.1 % of the patients aged younger than 12 years, AsthmaCritic pointed out that the fourth course of antibiotics in twelve months had been prescribed, and recommended the use of anti-inflammatory medication. However, although the Dutch guidelines recommend anti-inflammatory medication instead of repeated use of antibiotics, the physician may have good reasons to prescribe antibiotics. Other comments alert to clear deviations from the guidelines. For example, 11.0% of the patients aged 12 years and older, and 13.0% of the patients aged younger than 12 years received medication with a dose outside the recommended range; frequently, the physician had prescribed too low a dose.

Although most comments of AsthmaCritic are associated with specific recommendations (e.g., the recommendation to start long-acting bronchodilators), comments in the category *Alarming situations* (14.7 % of the patients aged 12 years

and older and 10.6 % of the patients aged younger than 12 years) and *Inconsistent Route* (9.7 % in patients aged 12 years and older and 3.6 % in patients aged younger than 12 years) only point out that the patient requires evaluation. AsthmaCritic, for example, detects decreasing peak flow measurements or increased consumption of bronchodilators and draws the attention of the physician to these trends; the clinical response is left to the physician.

Our study with actual electronic patient records from primary care practices shows that AsthmaCritic is both able to select asthma or COPD patient records and to generate patient-specific comments. The number of comments (on average, 3.4 per encounter) is considerable for daily practice. The acceptance of a decision-support system, however, not only depends on the number of comments, but also on the kind of comments generated and the way feedback is presented. If physicians' behaviour will be influenced, it is not clear whether the number of comments will increase or decrease in response to that changed behaviour. On one hand, one can expect the number of generated comments to reduce because the physician may decide to follow the guidelines (e.g., change his dosing schemas or start prescribing non-antibiotic anti-inflammatory medication in the appropriate cases). On the other hand, the system may stimulate a more complete recording of medical data, thereby increasing the system's ability to generate (more specific) comments. The acceptance of these comments may differ from comments made before the change in behaviour. Field studies will be needed to assess these effects.

AsthmaCritic is developed to be part of physicians' working environment. Integration with daily practice is the key. In addition to being able to deliver patient-specific feedback, integration implies leaving the physician in control and – if available – using routinely recorded data. As we have argued, leaving the physician in control is required from a medical point of view. In addition, if a decision-support system has to fit daily practice, the physician should be able to control the system to match his or her available time and needs at any moment. AsthmaCritic, therefore, has to provide the physician with tools enabling him to execute such control. Using routinely recorded data prevents the physician from having to record data twice and prevents workflow interruptions. In a previous study we observed that routinely recorded data are sufficient for human reviewers to generate patient-specific feedback[30]. This study shows that a computer-based decision-support system can generate patient-specific

feedback based on routinely recorded data, thereby enabling the physician to reflect on the treatment for an individual patient based on current guidelines. Additional studies will have to assess the validity and usability of produced feedback and if AsthmaCritic is able to change physicians' behaviour with respect to diagnosis and treatment of asthma and COPD.

ACKNOWLEDGEMENTS

We want to thank the members of our national medical content board for their share in the development of the AsthmaCritic knowledge base: B. Bottema, MD, PhD, P. N. R. Dekhuyzen, MD, PhD, E. J. Duiverman, MD, PhD, E. E. M. van Essen, MD, PhD, Th. B. Voorn, MD, PhD, M. H. J. Vaessen, MD, and A. van der Kuy, PhD.

REFERENCES

1. Rutten-van Molken MP, Postma MJ, Joore MA, Van Genugten ML, Leidl R, Jager JC. Current and future medical costs of asthma a nd chronic obstructive pulmonary disease in The Netherlands. Respir Med 1999;93(11):779-87.
2. Bottema BJAM, Fabels EJ, Van Grunsven PM, Van Hensbergen W, Muris JWM, Van Schayck CP, et al. NHG Standaard CARA bij Volwassenen: Diagnostiek [Guidelines of the Dutch College of General Practitioners: Chronic Respiratory Diseases in Adults: Diagnostics]. Huisarts en Wetenschap 1992;35(11):430-6.
3. Dirksen WJ, Geyer RMM, De Haan M, Kolnaar BGM, Merkx JAM, Romeijnders ACM, et al. NHG Standaard Astma bij Kinderen [Guidelines of the Dutch College of General Practitioners: Asthma in Children]. Huisarts en Wetenschap 1992;35(9):355-62.
4. Waart van der MAC, Dekker FW, Nijhoff S, Thiadens HA, Van Weel C, Helder M, et al. NHG Standaard CARA bij Volwassenen: Behandeling. [Guidelines of the Dutch College of General Practitioners: Chronic Respiratory Diseases in Adults: Therapy]. Huisarts en Wetenschap 1992;35(11):437-43.
5. Geijer RMM, Thiadens HA, Smeele IJM, Zwan van der AAC, Sachs APE, Bottema BJAM, et al. NHG-Standaard COPD en astma bij volwassenen: Diagnostiek [Guidelines of the Dutch College of General Practitioners: Chronic Obstructive Respiratory Diseases and Asthma in Adults: Diagnostics]. Huisarts en Wetenschap 1997;40(9):415-28.
6. Geijer RMM, Hensbergen van W, Bottema BJAM, Schayck van CP, Sachs APE, Smeele IJM, et al. NHG-Standaard astma bij volwassenen: Behandeling [Guidelines of the Dutch College of General Practitioners: Asthma in Adults: Therapy]. Huisarts en wetenschap 1997;40(9):443-54.
7. Geijer RMM, Schayck van CP, Weel van C, Sachs APE, Zwan van der AAC, Bottema BJAM, et al. NHG-Standaard COPD: Behandeling [Guidelines of the Dutch College of General Practitioners: Chronic Obstructive Respiratory Diseases: Therapy]. Huisarts en wetenschap 1997;40(9):430-42.
8. Dirksen WJ, Geijer RMM, De Haan M, De Koning G, Flikweert S, Kolnaar BGM. NHG-standaard astma bij kinderen [Guidelines of the Dutch College of General Practitioners: Asthma in Children]. Huisarts en wetenschap 1998;41(3):130-43.
9. Wyatt J. Uses and sources of medical knowledge. Lancet 1991;338:1368-72.
10. Haines A, Jones R. Implementing findings of research [see comments]. British Medical Journal 1994;308(6942):1488-92.
11. Freemantle N, Grilli R, Grimshaw J, Oxman A. Implementing findings of medical research: the Cochrane Collaboration on Effective Professional Practice. Quality in Health Care 1995;4(1):45-7.
12. Field MJ, Lohr KN. Clinical practice guidelines: directions for a new program: Institute of Medicine; 1990.

13. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't [editorial] [see comments]. British Medical Journal 1996;312(7023):71-2.

14. McColl A, Smith H, White P, Field J. General practitioners' perceptions of the route to evidence-based medicine: a questionnaire survey. British Medical Journal 1998;316:361-5.

15. Moulding N, Fahy N, Foong LH, Yeoh J, Silagy C, Weller D. A systematic review of the current status of evidence-based medicine and its potential application to Australian general Practice: Department of general practice, Flinders University of South Australia; 1997 January.

16. Woolf SH, Grol R, Eccles M, Grimshaw J. Potential benefits, limitations, and harms of clinical guidelines. British Medical Journal 1999;318:527-30.

17. Grimshaw J, Freemantle N, Wallace S, Russell I, Hurwitz B, Watt I, et al. Developing and implementing clinical practice guidelines. Quality in Health Care 1995;4:55-64.

18. Lomas J, Anderson GM, Domnick-Pierre K, Vayda E, Enkin MW, Hannah WJ. Do practice guidelines guide practice? The effect of a consensus statement on the practice of physicians. New England Journal of Medicine 1989;321(19):1306-11.

19. Kassirer JP, MD. A report card on computer-assisted diagnosis-the grade: C. New England Journal of Medicine 1994;330(25):1824-5.

20. Elson RB, MD, Connelly DP, MD, PhD. Computerized decision-support systems in primary care. Primary Care 1995;22(2):365-84.

21. Hunt DL, MD, Haynes RB, MD, PhD, Hanna SE, MA, PhD, Smith K. Effects of computer-based clinical decision support systems on physician performance and patient outcomes: A systematic review. Journal of the American Medical Association 1998;280:1339-46.

22. McDonald CJ, Wilson GA, McCabe GJJ. Physician response to computer reminders. Journal of the American Medical Association 1980;244:1579-81.

23. Linnarsson R. Decision support for drug prescription integrated with computer-based patient records in primary care. Medical Informatics 1993;18(2):131-42.

24. Miller RA, MD. Medical diagnostic decision support systems–past, present, and future: A threaded bibliography and brief commentary. Journal of the American Medical Informatics Association 1994;1:8-27.

25. Lei van der J, MD, PhD, Duisterhout JS, MSc, Westerhof HP, MD, Does van der E, MD, PhD, Cromme PVM, MD, PhD, Boon WM, MS, et al. The introduction of computer-based patient records in The Netherlands. Annals of Internal Medicine 1993;119(10):1036-41.

26. Mitchell E, Sullivan F. A descriptive feast but an evaluative famine: systematic review of published articles on primary care computing during 1980-97. British Medical Journal 2001;322(7281):279-82.

27. Boersma JJ. ICPC: International classification of primary care: short titels and Dutch subtitels. Utrecht: Nederlands Huisartsen Genootschap; 1995.

28. WHO International working group for drug statistics methodology. Anatomical therapeutic chemical (ATC) classification and defined daily dose

(DDD) for pharmaceuticals and vitamins. In. Oslo, Norway: WHO collaborating centre for drug statistics methodology; 1996.

29. Vlug AE, Lei van der J, Mosseveld BMT, Wijk van MAM, Linden van der PD, Sturkenboom MCJM, et al. Postmarketing surveillance based on electronic patient records: The IPCI project. Methods of Information in Medicine 1999;38:339-44.

30. Kuilboer MM, Lei van der J, Jongste de J, Overbeek S, Ponsioen B, Bemmel van JH. Exploring the role of an integrated critiquing system: A simulation. Journal of the American Medical Informatics Association 1998(5):194-202.

31. Gijsen R, Verkleij H, Dijksterhuis PH, Lisdonk van de EH, Metsemakers JFM, Velden van der J. Ziektespecifieke vergelijking van de geregistreerde morbiditeit in vier huisartsenregistraties: een analyse ten behoeve van VTV-1997 [Disease-specific comparison of recorded morbidity in four general practitioner practices: an analysis on behalf of the Public Health Status and Forecast - 1997]: Rijksinstituut voor Volksgezondheid en Milieu Bilthoven; 1997 August. Report No.: 431501017.

32. Registratienet Huisartsenpraktijken (Data network general practitioner practices). Gezondheidsproblemen en Diagnosen in de huisartsenpraktijk [Health care problems and diagnosis in general practitioner practices]: Rijksuniversiteit Limburg, Maastricht; 1993 01/03/93. Report No.: 1.

33. Waal de M. RNUH-LEO Basisrapport VI: LUMC Leiden, The Netherlands; 1998. Report No.: 6.

34. Barry PW, Fouroux B, Pedersen S, O'Callaghan C. Nebulizers in childhood. European Respiratory Journal 2000;10:527-35.

5

COMPUTERIZED CRITIQUING INTEGRATED WITH DAILY CLINICAL PRACTICE AFFECTS PHYSICIANS' BEHAVIOUR

A RANDOMISED CLINICAL TRIAL WITH ASTHMACRITIC

Submitted for publication

Manon M. Kuilboer
Mees Mosseveld
Marc A. M. van Wijk
Emiel van der Does
Johan C. de Jongste
Shelley E. Overbeek
Ben Ponsioen
Johan van der Lei

ABSTRACT

BACKGROUND

The quality of Asthma and COPD (chronic obstructive pulmonary disease) treatment is below current medical standards, causing under-treatment and unnecessary high health-care expenditure. Guidelines have been developed to support physicians providing up-to-date care. Computer-based decision-support systems (CDSSs) may help the implementation of guidelines because they have the potential to influence physician behaviour. We developed AsthmaCritic, a non-inquisitive critiquing system, integrated with general practitioners' electronic medical record. The system is based on the guidelines for asthma and COPD of the Dutch College of General Practitioners.

OBJECTIVE

To assess the effect of AsthmaCritic on monitoring and treatment of asthma and COPD by Dutch general practitioners in daily practice.

DESIGN

Randomised clinical trial.

SETTING

Primary care.

PARTICIPANTS

32 practices (40 Dutch general practitioners) using electronic patient records

INTERVENTIONS

Practices were randomised to an intervention group that was enabled to use AsthmaCritic or to a control group that continued working as usual.

MAIN OUTCOME MEASURES

Average number of contacts, FEV_1 (Forced expiratory volume), and peak flow measurements per patient per practice; average number of antihistamine, cromoglycate, deptropine, and oral bronchodilator prescriptions per patient per practice.

RESULTS

The number of contacts increased in the age group 12-39 years. The number of FEV_1, peak flow measurements, and the ratio of coded measurements increased, whereas the number of cromoglycate prescriptions decreased in the age group 12-39 years.

CONCLUSIONS

Our study shows that the guideline-based critiquing system AsthmaCritic changed physicians' monitoring and, to a lesser extent, treatment behaviour. In addition, the physicians changed their data recording habits.

INTRODUCTION

The quality of asthma and COPD treatment is below current medical standards[1-3]. It has proven hard for health-care professionals to keep up with rapidly changing insights into the diagnosis and treatment of asthma and COPD. To encourage the application of evidence based medicine in daily practice, professional health-care organisations developed guidelines to provide physicians with a summary of large volumes of clinical evidence and a related set of practical recommendations[4-7]. In the Netherlands, for example, over 80 guidelines exist, each comprising 2 to 4 pages - a paper-based hurdle for quick integration of new medical knowledge[8]. Internationally, the implementation of guidelines has also been disappointingly slow[9]. Despite the recommendation to use inhaled corticosteroids in patients with moderate and severe asthma, recent studies showed considerable under-use of inhaled corticosteroids by these patients[10]. Also, insufficient monitoring of patients' lung function was demonstrated[10]. Under-treatment and under-use of monitoring may lead to high health-care expenditure and sub-optimal patient care[11]. Computer-based decision-support systems (CDSSs) have been advocated to support the implementation of guidelines because they have the potential to influence physician behaviour[12]. In the Netherlands, the majority of the general practitioners replaced their paper-based patient record with an electronic patient record. They record patient data themselves, during the patient encounter. The Dutch infrastructure creates an opportunity to evaluate a non-inquisitive CDSS (i.e., a CDSS that does not interrupt the physician for additional data entry specifically for the CDSS). Given the known backlog of the implementation of asthma/COPD guidelines we developed AsthmaCritic[13, 14]. AsthmaCritic is based on the asthma/COPD guidelines issued by the Dutch College of General Practitioners[4-7]. If the system is able to influence physicians' behaviour, guideline-based recommendations for monitoring and treatment of asthma and COPD may be introduced more efficiently and thus improve health-care. We performed a randomised trial to assess the effect of AsthmaCritic on monitoring and treatment of asthma and COPD by Dutch general practitioners.

METHODS

INTERVENTION

To study the effect of critiquing systems integrated in daily practice we developed AsthmaCritic, a decision-support system that provides the general practitioner with patient-specific feedback on monitoring and treatment of patients with asthma or

COPD. To generate such feedback, the system uses all retrospective data[1] from the physician's information system only – no specific data entry is required. The physician does not need to activate the program - as soon as all data have been recorded feedback is automatically generated whenever specific data trigger the system[2]. The physician is notified of available feedback right away. He or she may choose to interrupt the analysis causing the program to run the analysis in the background. The physician can inspect the results later. In addition, if comments are generated in background mode, patient-specific messages are attached to the medical record, notifying the practitioner of available feedback. The physician can review the generated feedback the next time he opens the record, or anytime upon his own request. He may choose to ignore the feedback.

AsthmaCritic offers feedback as a list of comments each containing more specific information, more elaborated information, the reason why the comment had been generated, the comments' resource, as well as access to a structured form of the Dutch Guidelines. The physician is free to choose which information he prefers to read, depending on available time and knowledge. AsthmaCritic keeps a log of its use.

AsthmaCritic has predominantly been based on the asthma and COPD guidelines of the Dutch College of General Practitioners[17-19] and has been built over a period of 3 years under guidance of a *medical content board* (two local general practitioners and two pulmonologists, and seven national specialists in asthma and COPD).

PARTICIPANTS

In February 1998, all practices in the region of Delft, the Netherlands, were invited to participate in the study. All practices that had replaced the paper-based medical record with an electronic medical record of the information system ELIAS© and used the computer during patient encounters were eligible for the study.

1 Symptoms, diagnosis, measurements, referrals, tags, and problem list.
2 Triggers used by AsthmaCritic: ICPC codes (' International Classification of Primary Care – coding system for Diagnosis, symptoms and procedures'[15]) for asthma (R96), chronic bronchitis (R91), emphysema (R95), other chronic pulmonary diseases (R83.4), and the ATC code for prescriptions (the Anatomical, Therapeutic and Chemical (ATC) coding system of the World Health Organisation[16]) used in the treatment of asthma or COPD (R03).

RANDOMISATION

To evaluate the effect of AsthmaCritic, practices were stratified for single-handed or group practice and subsequently randomised either to the control group or to the AsthmaCritic group[20, 21]. To avoid a possible seasonal influence caused by differences in start of study participation (total installation period 2 months; July through August 1998), each practice of the AsthmaCritic group was matched with a practice of the control group.

PROTOCOL

In the Netherlands, patients are registered with one general practitioner. Both for the retrospective baseline study and the prospective intervention study we anonymized and downloaded the complete electronic patient record of all registered patients. In addition, the AsthmaCritic group received a one-hour instruction, a manual, a description of the system's knowledge base, and a memo card with coding details and helpdesk phone numbers at installation. After one week, a second visit was scheduled for the AsthmaCritic group to answer any questions about the use of the system. No other contact was sought until the end of the study.

STUDY POPULATION

All patients who were registered during (part of) the study period with one of the participating general practitioners and who had at least one of the following codes in their electronic medical record were part of our study population; ICPC*[3]* codes for asthma (R96), chronic bronchitis (R91), emphysema (R95), other chronic pulmonary diseases (R83.4), or the ATC code*[4]* for prescriptions used in the treatment of asthma or COPD (R03).

STUDY PERIOD

The study period was divided into a five-month baseline period and a five-month intervention period.

OUTCOME PARAMETERS

In the 6 months prior to the study, the Dutch College of General Practitioners issued revised guidelines for the diagnosis and treatment of asthma and COPD[4-7]. These revised guidelines replaced guidelines dating from 1992 and 1993. Each revised guideline is preceded by a section detailing the major changes in the guideline

3 International Classification of Primary Care – coding system for diagnosis, symptoms and procedures[15].
4 The Anatomical, Therapeutic, and Chemical coding system of the World Health Organisation[16].

compared with the previous version of that guideline; the outcome parameters were primarily based on these changes and included:

- Contact frequency (the revised guidelines recommend increased contact frequency for new patients and patients whose medication is being changed);
- Peak flow measurements (the revised guidelines emphasize the use of peak flow measurements in the diagnostic phase of children and adults, and in the treatment phase of children and adults with asthma)
- FEV1 (the revised guidelines emphasize the importance of spirometry for COPD)
- Cromoglycate (the revised guidelines limit the use of cromoglycates to those children that do not tolerate inhaled corticosteroids or adults with allergic asthma only)
- Deptropine (the revised guidelines discourage the use of this drug, that used to be prescribed for children)
- Antihistamines (the revised guidelines discourage the use of these drugs, that used to be prescribed for adults with asthma and for children)
- Oral bronchodilators (for children, the revised guidelines recommend inhaled medication instead of oral medication).

We counted for all patients in our study population the average number of contacts, measurements, and prescriptions per patient per practice as defined in the next three paragraphs.

CONTACTS

A contact is defined as a physical or phone-based contact between a patient and the practice, excluding contacts generated by repeat prescriptions and incoming laboratory tests recorded by the practice assistant. Two contacts on one day were counted as one.

MEASUREMENTS

We counted all peak flow and FEV_1 measurements recorded in the electronic patient record of our study population; both the free text additions to the electronic patient record and the measurements recorded using the general practitioner information system's internal coding system.

PRESCRIPTIONS

Based on the ATC codes for prescriptions, we counted the average number of antihistamine, cromoglycate, deptropine, and oral bronchodilator prescriptions.

STATISTICAL ANALYSIS

Comparisons of baseline characteristics between the study groups were done with the Chi-square test for categorical variables and the Mann Whitney U test for continuously distributed variables. In addition, we also compared the outcome parameters for the baseline period by the Mann Whitney U test. The effect of AsthmaCritic was assessed by calculating the intra-practice changes in outcome parameters between the baseline and intervention period within age groups (0-11, 12-39, 40-59, and ≥ 60 years) that were based on age-specific guidelines (delta value). These delta values were compared between the control and AsthmaCritic group with a paired Wilcoxon signed rank test. All analysis were done with SPSS version 10.0.7. All statistical tests were two-sided and comparisons with an error probability smaller than 5% were considered statistically significant.

RESULTS

Of the 141 general practices in the Delft region, 32 practices agreed to participate and completed the study. Sixteen practices, also involving 20 general practitioners were assigned to the control group and sixteen practices, involving 20 general practitioners, were assigned to the AsthmaCritic group.

BASELINE DATA

At the start of the intervention period (the practices started between June 30th 1998 and August 22nd 1998), 78,926 patients were enrolled in the practices assigned to the control group. Of these patients, 41,867 (53.05 %) were insured by governmental insurance, 34,633 (51.1%) were men, and the average age was 36.7 years (SD 4.2 years). In the same period a total of 77,846 patients were enrolled in the practices assigned to AsthmaCritic; 41,929 (53.86 %) patients were insured by governmental insurance, 34,083 (50.4%) were men, and the average age was 38.4 years (SD 3.4 years). Table 1 shows the baseline characteristics of the practice populations at start of intervention. These characteristics were not statistically significantly different. Table 2 shows the baseline characteristics of the general practitioners at start of intervention. The two groups of general practitioners differed for the baseline characteristic age; the general practitioners in the AsthmaCritic group being, on average, about 3 years older. There was no significant correlation in the control group

between the age of the general practitioners and the outcome parameters (results not shown). The general practitioners did not differ for the baseline characteristics Continuous Medical Education (CME) credits, or the period since the latest CME course on asthma or COPD had been followed. The proportion of physicians with a special interest in asthma or COPD or in computers was comparable between the groups. In all age groups, in the baseline period, there were no significant differences in the eight outcome parameters between the two study groups (results not shown).

ASTHMACRITIC USE

Analysis of a record took 31.7 seconds on average (SD 17.6 seconds). The physicians waited for the analysis in, on average, 22% of the times an analysis was completed. The physicians reviewed 32% of the generated feedback at least once in the study period. Feedback for the participating physicians was generated 10,863 times; 10,532 with comments and 331 times with a message that no comments had been made. The median time spent by the physician reviewing generated feedback was 9 seconds (25^{th} percentile = 4 seconds, 75^{th} percentile = 48 seconds).

MONITORING

Table 3 shows the results for the average number of contacts, peak flow measurements, and FEV_1 measurements. In the AsthmaCritic group, the number of contacts increased more than in the control group. This difference was statistically significant in age group 12-39 (paired Wilcoxon signed rank test, $p = 0.034$). Figure 1 graphically illustrates the results for the average number of contacts.

In the AsthmaCritic group, the average number of peak flow measurements per patient per practice increased more than in the control group. This difference was statistically significant in age group 0-11 and 12-39 years ($p = 0.016$, $p = 0.020$, respectively). For the average number of FEV_1 measurements the increase was statistically significant for age group 0-11 ($p = 0.028$). See Figure 2 and 3 that graphically illustrate the results for the average number of peak flow and FEV_1 measurements.

The ratio of coded peak flow measurements over all peak flow measurements increased more in the AsthmaCritic group. This difference was statistically significant in age group 12-39 and 40-59 ($p = 0.004$, $p = 0.009$). For FEV_1 measurements, the same ratio increased in all age groups ($p = 0.046$, $p = 0.010$, $p = 0.010$, $p = 0.016$).

TREATMENT

For age group 12-39 years the average number of cromoglycate prescriptions decreased more in the AsthmaCritic group (paired Wilcoxon signed rank test p = 0.033). Deptropine, antihistamines, and oral bronchodilators did not show statistically significant changes. Figure 4 graphically illustrates the results for the average number of cromoglycate prescriptions. Table 4 presents an overview of the results.

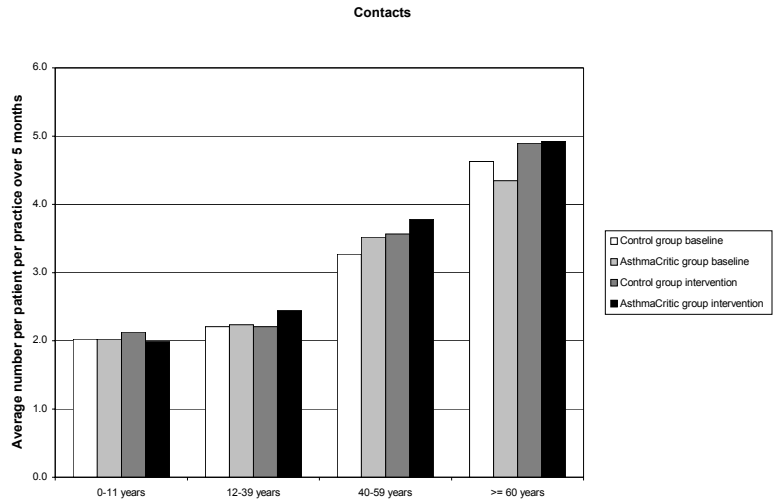


FIGURE 1.

AVERAGE NUMBER OF CONTACTS PER PATIENT PER PRACTICE BY AGE GROUP.

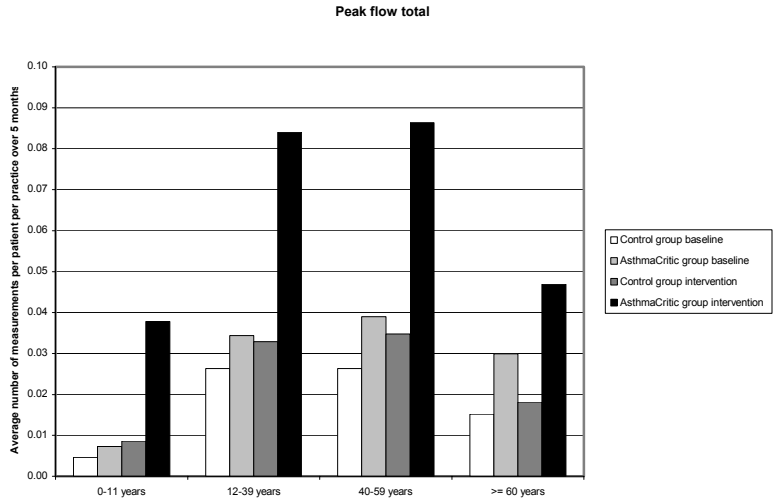


FIGURE 2.

AVERAGE NUMBER OF PEAK FLOW MEASUREMENTS PER PATIENT PER PRACTICE BY AGE GROUP.

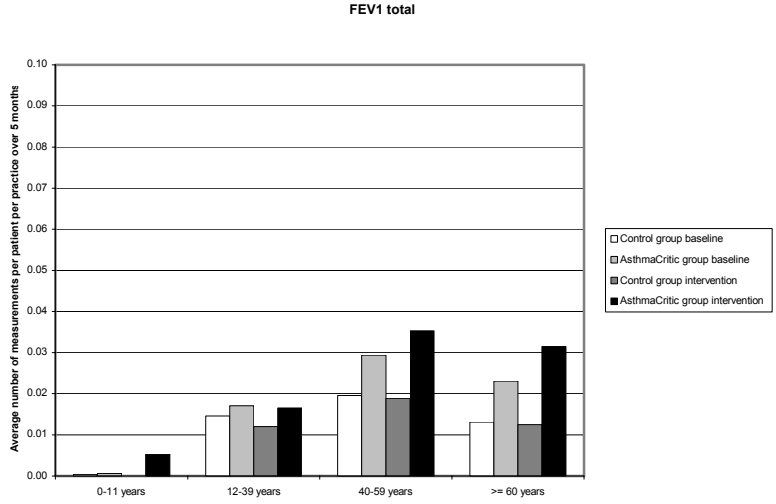


FIGURE 3.

AVERAGE NUMBER OF FEV_1 MEASUREMENTS PER PATIENT PER PRACTICE BY AGE GROUP.

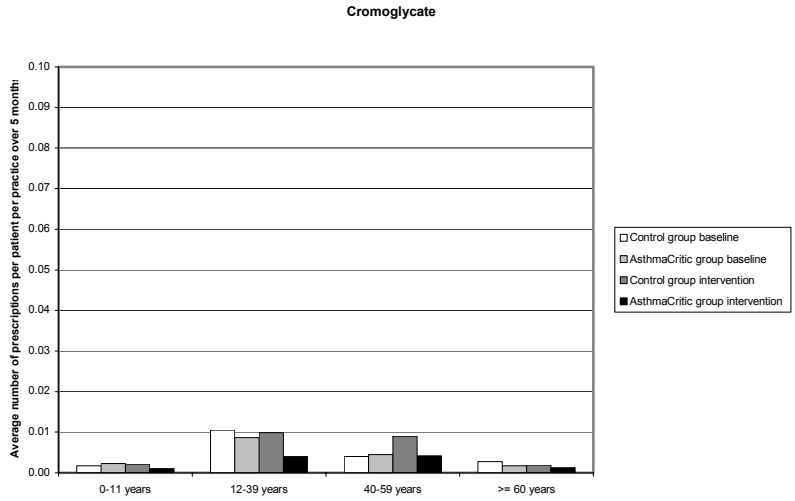


FIGURE 4.

AVERAGE NUMBER OF CROMOGLYCATE PRESCRIPTIONS PER PATIENT PER PRACTICE BY AGE GROUP.

DISCUSSION

We assessed the effect of AsthmaCritic on monitoring and treatment of asthma and COPD by Dutch general practitioners in a randomised controlled trial in daily practice (a five-month baseline period followed by a five-month intervention period). The system was integrated with the electronic patient records used by the practitioner during consultation. Based on routine data only, without additional data-entry effort required from the physician, feedback was presented during daily routine.

AsthmaCritic is based on the Dutch asthma and COPD guidelines, which had been revised just before our study started. The outcome parameters of our study were based on the description of the most significant changes in the new guidelines compared to the previous version of that guideline. The outcome parameters involved monitoring (frequency of contacts, of peak-flow measurements, and of FEV_1) and treatment (prescription of drug categories).

The use of AsthmaCritic was associated with an increase in the number of contacts (absolute increase with about 5-10%) and the number of times physicians assessed their patients' pulmonary function. The increase of the average number of peak flow measurements was larger (absolute increase with up to over 50%) than the increase on the average number of FEV_1 measurements (absolute effect smaller and variable). This difference is probably due to the difference in logistics for each of these measurements. Easy logistics are amongst the known facilitators for changes in practice management[22]. Peak flow measurements can be easily performed at the general practitioner's surgery, while getting an FEV_1 required a referral to the pulmonologist or a special laboratory facility. We believe that the observed increase of pulmonary function assessments is clinically relevant. Some investigators have emphasised the discrepancy between pulmonary disease symptoms and disease severity[23], and have shown in observational studies an under utilisation of pulmonary function assessment[3]. We conclude that AsthmaCritic changed the physicians' monitoring of asthma and COPD.

To evaluate AsthmaCritic's effect on treatment, we focussed on four drug categories that had their use curtailed by the new guidelines. Our study showed, in line with the recommendations of the revised guideline, a decrease of the number of cromoglycate prescriptions in the AsthmaCritic group. For the other drugs (deptropine, antihistamines, and oral bronchodilators), no changes were observed. Although the guidelines emphasized as major changes the fact that these drugs should no longer be prescribed, we observed that general practitioners already hardly used these drugs (for example, in the control group in 40-59 year-old patients cromoglycate was prescribed 4 times per 1000 patients over 5 months, which is 10 times per 1000 patients per year). Other investigators have shown that physicians may change their behaviour prior to the publication of revisions of guidelines[24].

In addition to the changes in monitoring and treatment, the physicians changed their data-recording habits, as can be seen by the increase in the ratio of coded measurements over the sum of measurements recorded as a free-text addition to the electronic patient record and the coded measurements. Availability of structured data is vital for CDSSs to produce patient-specific feedback. In a previous study, we showed that sufficient data -- a mix of structured data and free text -- were available in physicians' electronic patient record to generate feedback[25]. AsthmaCritic, as a non-

inquisitive system, does not force physicians to structure their data. In response to AsthmaCritic, physicians did record more data in a structured fashion, thus allowing the system to include more data in the analysis.

AsthmaCritic was used by the practitioners during their normal routine. After the consultation of an asthma/COPD patient, the system analysed the complete medical record including the data about the current consultation. The general practitioners were able to interrupt the analysis of a record ('one hit on any key'). An analysis took, on average, about 30 seconds. The general practitioners waited for the feedback in one-fifth of the cases; in the remaining cases they did not wait for the system to finish the analysis, but were alerted to generated feedback by a patient-specific message. A question that remains unanswered in our study is whether AsthmaCritic would have a larger impact on physicians' behaviour if the processing time would be reduced to just a few seconds. The answer may not be straightforward – physician attitude, information (over-)load, and effect fade-out are only a few of the factors that may influence the outcome[12, 13, 26]. Further research into the relation between system characteristics and effect on physician behaviour will be needed to answer this question.

AsthmaCritic generates feedback irrespective of the reason for encounter. As a result, AsthmaCritic may generate feedback even if the contact does not cover asthma- or COPD-related issues, possibly causing averse responses from the physician. Previous research, however, shows that many patients will visit their physician for their chronic pulmonary disease only when symptoms deteriorate[10]. Feedback irrespective of the reason for encounter may improve patient monitoring because it prompts the physician to assess the asthma/COPD status even when the patient comes for a different problem.

In this study, we focussed on guidelines for asthma/COPD. If additional guidelines are to be included in a critiquing system, a number of aspects need consideration. Firstly, professional organisations typically develop paper guidelines. A number of researchers have emphasized that translating paper guidelines into electronic decision-support systems is a time-consuming effort and may show inconsistencies and ambiguities[27]. We believe that wide-scale use of decision-support systems as a means to implement guidelines will require professional organisations to anticipate the use of decision-support systems in all stages of guidelines development.

Secondly, if feedback will be generated on many different patient categories, the result may be an abundance of alerts. Too many alerts may cause the feedback to be ignored. Further research is needed to study possible approaches to ensure optimal impact without causing information overload.

In conclusion, we showed that AsthmaCritic changed physician behavior and influenced monitoring and, to a lesser extent, treatment of asthma and COPD in general practice.

ACKNOWLEDGEMENTS

We thank the general practitioners who took part in this study for their time, enthusiasm and criticisms - J.M. Baks, J.P. Bijl, C.M.J. Bonekamp, C.S. Bos, H. Breedveldt-Boer, J. Breugem, J.A. Brienen, P.J.A. Bucx, P. de Blooy, R.H. Dupuis, J.A.J. Garretsen, R. Glotzbach, C. Jansen, C.H.F. Jonker, W. Kamermans, L.E.M. Kleipool, A. Klokke, V.J. Knippenberg, P.E.H. Kromdijk, P.J.Th.M Meijs, K.P. Nederlof, J.B.M. Nijkamp, J. Oosthoek, M.A. Plasmans, L. Redel, A.R.N. Rijckevorsel, F.J.N. Rijkee, G. Slagter, H.S. Spijker, A.F.M. Spreeuw, B. Sprij, F.C.M. Touw, W. van Donselaar, M.J. van Dijk, R.D.W. van Bentveld, L.J. van Loon, S.J. van 't Lindenhout, R. van Stijn, R.T.M. Veldkamp, H.W. Visser, H.J.P. Vos, W. Wierema.

We also thank the members of the national medical content board - B. Bottema, MD, PhD, P.N.R. Dekhuyzen, MD, PhD, E.J. Duiverman, MD, PhD, E.E.M. van Essen, MD, PhD, Th. B. Voorn, MD, PhD, M.H.J. Vaessen, MD, and A. van der Kuy, PhD, for their share in the development of the AsthmaCritic knowledge base.

Finally, we thank the following three statisticians for their advice with the statistical analysis: J.P.C. Dieleman, PhD, Prof. Th. Stijnen, PhD and M.C.J.M. Sturkenboom, PhD.

TABLE 1.

BASELINE CHARACTERISTICS OF PRACTICES AT START OF INTERVENTION

CHARACTERISTICS	CONTROL GROUP (N= 16)				ASTHMACRITIC GROUP (N= 16)				P VALUE †
	MEAN	MEDIAN	25TH	75TH	MEAN	MEDIAN	25TH	75TH	
Enrolled patients, n	4933	4359	3689	6866	4865	4686	3734	6009	.880
Patient Age, y	36.7	37.4	34.2	39.6	38.4	37.9	36.6	38.8	.498
Male patients, %	50.8	50	49.4	50.9	50.4	49.8	49.2	51.4	.678
Patients insured through government insurance, %	53.2	52.4	50.7	60.6	54	54.8	49.3	59.2	.970
Asthma/COPD patients,%	9.8	9.3	6.8	10.5	11.1	11.1	9.4	12.6	.083

† 2-sided Mann Whitney U test on practice means

TABLE 2.

BASELINE CHARACTERISTICS OF THE GENERAL PRACTITIONERS AT START OF INTERVENTION.

CHARACTERISTICS	CONTROL GROUP				ASTHMACRITIC GROUP				P VALUE†
	MEAN	MEDIAN	25TH	75TH	MEAN	MEDIAN	25TH	75TH	
age at study start y (n=20, 20)	43.1	42.5	39.3	46.8	46.5	47.0	42.8	49.8	.026 *
Time since latest CME on asthma/COPD y (n=19, 20)	0.95	1.00	1.00	1.00	0.95	1.00	0.00	1.00	.519 *
CME credits in 1997 (n=16, 16)	62.7	52.3	31.6	77.3	52.8	48.5	30.3	75.0	.749 *
CME credits in 1998 (n=16, 16)	62.9	64.3	33.1	82.9	58.5	64.0	35.1	70.5	.720 *
Special interest in Asthma/COPD (n=20, 20)	50 %				30 %				.20 ‡
Special interest in computers, (n=20, 20)	20 %				35 %				.29 ‡

* Mann-Whitney U test ‡ chi-square test for independence

TABLE 3

AVERAGES OF THE AVERAGE NUMBER PER PATIENT PER PRACTICE, BY STUDY PERIOD, BY STUDY GROUP, BY AGE GROUP, THE MEDIAN AND QUARTILES OF PAIRED DIFFERENCES OF DELTA VALUES[5] FOR NUMBER OF CONTACTS, PEAK FLOW MEASUREMENTS, AND FEV1 MEASUREMENTS A POSITIVE VALUE OF THE MEDIAN DELTA VALUE DENOTES A RELATIVE INCREASE IN THE ASTHMACRITIC GROUP.

MEASUREMENT AGE GROUP	BASELINE PERIOD		INTERVENTION PERIOD		MEDIAN OF PAIRED DIFFERENCES OF DELTA VALUES	QUARTILES		
	CONTROL GROUP	ASTHMA-CRITIC GROUP	CONTROL GROUP	ASTHMA-CRITIC GROUP		25TH	75TH	P-VALUE[6]
Contacts								
0-11	2.023	2.025	2.128	1.988	- 0.164	- 0.447	+ 0.181	0.255
12-39	2.206	2.237	2.210	2.443	+ 0.154	- 0.076	+ 0.468	0.034
40-59	3.270	3.517	3.564	3.784	+ 0.068	- 0.263	+ 0.150	0.756
≥60	4.628	4.344	4.898	4.929	+ 0.257	- 0.214	+ 0.685	0.134
Peak flow total[7]								
0-11	0.0046	0.0073	0.0086	0.0378	+ 0.020	0.000	+ 0.042	0.016
12-39	0.0263	0.0344	0.0329	0.0840	+ 0.029	- 0.004	+ 0.087	0.020
40-59	0.0263	0.0390	0.0348	0.0864	+ 0.028	- 0.029	+ 0.078	0.096
≥60	0.0152	0.0299	0.0182	0.0469	+ 0.005	- 0.004	+ 0.022	0.133
Peak flow ratio[8]								
0-11	0.0313	0.2500	0.0625	0.6115	+ 0.000	0.000	+ 1.000	0.071
12-39	0.2025	0.2646	0.1574	0.6643	+ 0.402	0.000	+ 0.901	0.004
40-59	0.1875	0.2900	0.1581	0.6973	+ 0.181	0.000	+ 0.950	0.009
≥60	0.1842	0.3094	0.1736	0.5546	+ 0.000	0.000	+ 0.860	0.108
FEV1 total								
0-11	0.0004	0.0006	0.0000	0.0053	+ 0.005	0.000	+ 0.007	0.028
12-39	0.0146	0.0171	0.0120	0.0165	+ 0.005	0.000	+ 0.015	0.062
40-59	0.0196	0.0294	0.0189	0.0353	+ 0.004	- 0.010	+ 0.026	0.422
≥60	0.0131	0.0231	0.0125	0.0315	0.000	- 0.003	+ 0.015	0.260
FEV1 ratio								
0-11	0.0625	0.0625	0.0000	0.2500	+ 0.000	0.000	+ 0.750	0.046
12-39	0.1823	0.0604	0.1736	0.4375	+ 0.056	0.000	+ 1.000	0.010
40-59	0.1875	0.1528	0.1844	0.6083	+ 0.250	0.000	+ 1.000	0.010
≥60	0.1217	0.1838	0.1996	0.4943	+ 0.000	0.000	+ 0.977	0.016

[5] 'Delta value' denotes the difference between the average number per patient per practice in the Intervention period (5-months) and the Baseline period (5 months).
[6] Paired Wilcoxon signed ranks test
[7] 'Total' denotes the number of measurements recorded as free text additions to the medical record plus measurements recorded using the information system's internal coding system.
[8] 'Ratio' denotes the ratio of measurements recorded using the information system's internal coding system over 'total'.

TABLE 4

AVERAGES OF THE AVERAGE NUMBER PER PATIENT PER PRACTICE, BY STUDY PERIOD, BY STUDY GROUP, BY AGE GROUP AND MEDIAN OF PAIRED DIFFERENCES OF DELTA VALUES[9] FOR AVERAGE NUMBER OF PRESCRIPTIONS BY PRESCRIPTION CATEGORY. A POSITIVE VALUE OF THE MEDIAN DELTA VALUE DENOTES A RELATIVE INCREASE IN THE ASTHMACRITIC GROUP.

MEASUREMENT AGE GROUP	BASELINE PERIOD		INTERVENTION PERIOD		MEDIAN OF PAIRED DIFFERENCES OF DELTA VALUES	QUARTILES		P-VALUE[10]
	CONTROL GROUP	ASTHMA-CRITIC GROUP	CONTROL GROUP	ASTHMA-CRITIC GROUP		25TH	75TH	
Antihistamines								
0-11	0.0087	0.0168	0.0078	0.0192	0.000	- 0.026	+ 0.014	0.875
12-39	0.0011	0.0013	0	0.0002	0.000	- 0.000	0.000	0.500
40-59	0.0004	0.0023	0.0015	0.0022	- 0.004	0.000	0.000	0.080
≥60	0	0.0004	0	0	0.000	0.000	0.000	0.317
Cromoglycate								
0-11	0.0017	0.0023	0.0021	0.0010	0.000	0.000	0.000	0.144
12-39	0.0105	0.0087	0.0099	0.0041	- 0.004	- 0.009	0.000	0.033
40-59	0.0040	0.0045	0.0090	0.0042	0.000	- 0.008	0.000	0.051
≥60	0.0028	0.0017	0.0019	0.0013	0.000	0.000	0.000	0.893
Deptropine								
0-11	0.0084	0.0278	0.0094	0.0260	- 0.003	- 0.020	+ 0.014	0.753
12-39	0	0	0	0	0.000	0.000	0.000	0
40-59	0	0	0	0	0.000	0.000	0.000	0
≥60	0	0	0	0	0.000	0.000	0.000	0
Oral bronchodilators								
0-11	0.0097	0.0133	0.0105	0.0093	+ 0.001	- 0.007	+ 0.008	0.807
12-39	0	0	0.0005	0.0003	0.000	0.000	0.000	0.655
40-59	0.0028	0	0.0026	0.0028	0.000	0.000	0.000	0.121
≥60	0.0011	0.0008	0.0031	0.0013	0.000	0.000	0.000	0.225

[9] 'Delta value' denotes the difference between the average number per patient per practice in the Intervention period (5-months) and the Baseline period (5 months).
[10] Paired Wilcoxon signed ranks test

REFERENCES

1. Taylor DM, Auble TE, Calhoun WJ, Mosesso VN, Jr. Current outpatient management of asthma shows poor compliance with International Consensus Guidelines. Chest 1999;116(6):1638-45.
2. Legorreta AP, Liu X, Zaher CA, Jatulis DE. Variation in managing asthma: experience at the medical group level in California. Am J Manag Care 2000;6(4):445-53.
3. Rabe KF, Vermeire PA, Soriano JB, Maier WC. Clinical management of asthma in 1999: the Asthma Insights and Reality in Europe (AIRE) study. Eur Respir J 2000;16(5):802-7.
4. Geijer RMM, Thiadens HA, Smeele IJM, Zwan van der AAC, Sachs APE, Bottema BJAM, et al. NHG-Standaard COPD en astma bij volwassenen: Diagnostiek [Guidelines of the Dutch College of General Practitioners: Chronic Obstructive Respiratory Diseases and Asthma in Adults: Diagnostics]. Huisarts en Wetenschap 1997;40(9):415-28.
5. Geijer RMM, Hensbergen van W, Bottema BJAM, Schayck van CP, Sachs APE, Smeele IJM, et al. NHG-Standaard astma bij volwassenen: Behandeling [Guidelines of the Dutch College of General Practitioners: Asthma in Adults: Therapy]. Huisarts en wetenschap 1997;40(9):443-54.
6. Geijer RMM, Schayck van CP, Weel van C, Sachs APE, Zwan van der AAC, Bottema BJAM, et al. NHG-Standaard COPD: Behandeling [Guidelines of the Dutch College of General Practitioners: Chronic Obstructive Respiratory Diseases: Therapy]. Huisarts en wetenschap 1997;40(9):430-42.
7. Dirksen WJ, Geijer RMM, De Haan M, De Koning G, Flikweert S, Kolnaar BGM. NHG-standaard astma bij kinderen [Guidelines of the Dutch College of General Practitioners: Asthma in Children]. Huisarts en wetenschap 1998;41(3):130-43.
8. Smeele IJ, Van Schayck CP, Van Den Bosch WJ, Van Den Hoogen HJ, Muris JW, Grol RP. [Discrepancy between the guidelines and practice by family physicians in treating adults with an exacerbation of asthma or chronic obstructive pulmonary disease]. Ned Tijdschr Geneeskd 1998;142(42):2304-8.
9. Crim C. Clinical practice guidelines vs actual clinical practice. The asthma paradigm. Chest 2000;118:62S-64S.
10. Vermeire PA, Rabe KF, Soriano JB, Maier WC. Asthma control and differences in management practices across seven European countries. Respir Med 2002;96(3):142-9.
11. Stoloff S. Current asthma management: the performance gap and economic consequences. Am J Manag Care 2000;6(17 Suppl):S918-25; discussion S925-9.
12. Morris AH. Developing and implementing computerized protocols for standardization of clinical decisions. Ann Intern Med 2000;132(5):373-83.
13. Elson RB, Faughan JG, Connelly DP. An industrial process view of information delivery to support clinical decision making: Implications for

systems design and process measures. Journal of the American Medical Informatics Association 1997;4(4):266-278.
14. Kuilboer M, van Wijk M, Mosseveld B, van der Does E, Ponsioen B, de Jongste J, et al. Feasibility of AsthmaCritic, a decision-support system for asthma and COPD, which generates patient-specific feedback on routinely recorded data in general practice. Family Practice 2002.
15. Boersma JJ. ICPC: International classification of primary care: short titels and Dutch subtitels. Utrecht: Nederlands Huisartsen Genootschap; 1995.
16. WHO International working group for drug statistics methodology. Anatomical therapeutic chemical (ATC) classification and defined daily dose (DDD) for pharmaceuticals and vitamins. In. Oslo, Norway: WHO collaborating centre for drug statistics methodology; 1996.
17. Dirksen WJ, Geyer RMM, De Haan M, Kolnaar BGM, Merkx JAM, Romeijnders ACM, et al. NHG Standaard Astma bij Kinderen [Guidelines of the Dutch College of General Practitioners: Asthma in Children]. Huisarts en Wetenschap 1992;35(9):355-62.
18. Bottema BJAM, Fabels EJ, Van Grunsven PM, Van Hensbergen W, Muris JWM, Van Schayck CP, et al. NHG Standaard CARA bij Volwassenen: Diagnostiek [Guidelines of the Dutch College of General Practitioners: Chronic Respiratory Diseases in Adults: Diagnostics]. Huisarts en Wetenschap 1992;35(11):430-6.
19. Waart van der MAC, Dekker FW, Nijhoff S, Thiadens HA, Van Weel C, Helder M, et al. NHG Standaard CARA bij Volwassenen: Behandeling. [Guidelines of the Dutch College of General Practitioners: Chronic Respiratory Diseases in Adults: Therapy]. Huisarts en Wetenschap 1992;35(11):437-43.
20. Kerry SM, Bland JM. Trials which randomize practices I: how should they be analysed? Fam Pract 1998;15(1):80-3.
21. Eccles M, Grimshaw J, Steen N, Parkin D, Purves I, McColl E, et al. The design and analysis of a randomized controlled trial to evaluate computerized decision support in primary care: the COGENT study. Fam Pract 2000;17(2):180-6.
22. Grol R, Dalhuijsen J, Thomas S, Veld C, Rutten G, Mokkink H. Attributes of clinical guidelines that influence use of guidelines in general practice: observational study. Bmj 1998;317(7162):858-61.
23. Osborne ML, Vollmer WM, Pedula KL, Wilkins J, Buist AS, O'Hollaren M. Lack of correlation of symptoms with specialist-assessed long-term asthma severity. Chest 1999;115(1):85-91.
24. van Wijk MA, van Der Lei J, Mosseveld M, Bohnen AM, van Bemmel JH. Compliance of General Practitioners with a Guideline-based Decision Support System for Ordering Blood Tests. Clin Chem 2002;48(1):55-60.
25. Kuilboer MM, Lei van der J, Jongste de J, Overbeek S, Ponsioen B, Bemmel van JH. Exploring the role of an integrated critiquing system: A simulation. Journal of the American Medical Informatics Association 1998(5):194-202.
26. Shiffman RN, Liaw Y, Brandt CA, Corb GJ. Computer-based guideline implementation systems: a systematic review of functionality and effectiveness. J Am Med Inform Assoc 1999;6(2):104-14.

27. Shiffman RN, MD, MCIS. Towards effective implementation of a pediatric asthma guideline: integration of decision support and clinical workflow support. In: Ozbolt JG, PhD, RN, editor. Proceeding of the Annual Symposium on Computer Application in Medical Care; 1994; Washington, DC, USA: Hanley & Belfus, Inc.; 1994. p. 797-801.

6

SUMMARY, DISCUSSION AND FUTURE RESEARCH

INTRODUCTION

The goal of this research was to further explore the potential of critiquing systems as tools to support physicians in performing health care according to current medical insight. This study focused on the evaluation of the feasibility and the effect on general practitioners' behavior of a critiquing system integrated with an electronic patient record in daily practice. We simulated, built, tested, and evaluated a critiquing system in the domain of asthma and chronic obstructive pulmonary disease (COPD). The next paragraphs summarize each of these steps and reflect on insights gained.

SIMULATING A CRITIQUING SYSTEM

Building and evaluating prototypes is a time-consuming effort to gain insight into design issues. We, therefore, started with a simulation study, which enabled us to identify some of the core aspects of our system design. A critiquing system requires electronically recorded patient data to generate feedback. Even though the amount and content of routinely recorded data in general practitioners' electronic patient record may be sufficient to fulfill physicians' needs (a record – 'reminder' of past events), these data may very well be insufficient to fulfill a critiquer's needs.

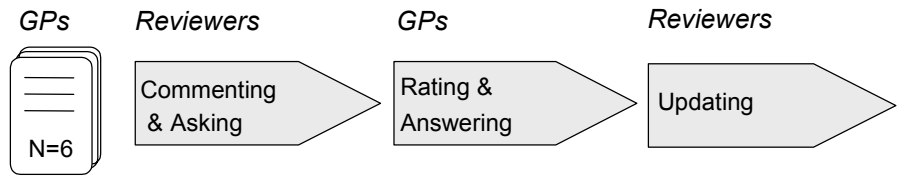


FIGURE 1.

THE THREE CONSECUTIVE STEPS IN THE SIMULATION STUDY. FOUR REVIEWERS ANALYZED SIX MEDICAL RECORDS. THE REVIEWERS GENERATED COMMENTS AND REQUESTED FURTHER INFORMATION WHEN NEEDED ('MISSED INFORMATION'). THE GENERAL PRACTITIONERS RATED THESE COMMENTS AND PROVIDED THE MISSING INFORMATION. WHEN INFORMATION WAS NOT AVAILABLE, THEY WERE ASKED TO EXPLAIN WHY. FINALLY, THE REVIEWERS UPDATED THEIR COMMENTS, TAKING THE ADDITIONAL INFORMATION INTO ACCOUNT.

In addition, time is limited and interruptions of physicians' normal routine to request additional data needed for the critiquing process easily experienced as annoying, even if these interruptions are meant to improve support. Therefore, we wanted to know whether the amount of patient data in the electronic medical record of general practitioners would suffice for the generation of critiquing comments. If the amount of patient data would be insufficient, we wanted to know what information would be

missing and how important this missing information would be for the generation of critiquing comments.

In the simulation study, described in Chapter 2, we asked four reviewers (two general practitioners and two specialists with a special interest in asthma or COPD) to play the role of the computer and generate critiquing comments on electronic medical records of patients with asthma or COPD. We asked three general practitioners to play the role of the users, assess these comments, and provide information being missed from the records by the reviewers when requested. Finally, we asked the four reviewers to reevaluate their own critiquing comments after the missing information had been provided by the three general practitioners. Figure 1 shows the different steps in the simulation study.

The study showed that different kinds of critiquing comments could be generated, and that much of the information that had been missed by the reviewers became available upon request. The reviewers left three-quarters of their comments unchanged after requested information had been made available, therefore, we decided for a non-inquisitive design.

THUS WE CONCLUDED THAT USING EPRS AS THE SINGLE DATA SOURCE FOR A CRITIQUING SYSTEM FOR ASTHMA OR COPD WAS FEASIBLE AND WE CHOSE TO BUILD A NON-INQUISITIVE SYSTEM.

In the simulation study, we asked the general practitioners why data being missed by reviewers from the records had not been recorded. The general practitioners most frequently mentioned that information had not been recorded explicitly or that it had been recorded elsewhere in the record (not necessarily accessible for a computer program). These answers illustrate the tension between physicians' data-recording needs and critiquers' data recording needs. To stimulate physicians recording data more explicitly and structured, tools for structured data recording need to be improved. The development of systems for structured data entry remains subject of active research – the challenge being how to combine complexity with clarity and ease of use[1].

DESIGNING AN INTEGRATED SYSTEM

Having decided for a non-inquisitive design, in Chapter 3 we further describe the design choices we made for the critiquing system AsthmaCritic, and reflect on issues underlying our choices with respect to system acceptance. Our precondition was to design an integrated system for the general-practice working environment. With 'integrated' we denote a system that receives its data from the general practitioners' information system as well as a system that aims to fit into general practitioners' daily practice.

To implement AsthmaCritic we extended the generic critiquing model from Van der Lei to fulfill the needs of a critiquing system aimed to function in physicians' daily practice[2]. The generic critiquing model supports the integration with an electronic medical record at data level by using events as the bridge between the electronic medical record and the critiquing system. Van der Lei built a prototype – HyperCritic - to evaluate the model in the domain of hypertension in a laboratory situation. The domain of the chronic diseases asthma and COPD is more complex than the domain of the risk factor hypertension. For example, to assess the severity of a patient's deterioration, the system needs to assess the frequency of symptoms, prescriptions, and pulmonary function measurements all together. Therefore, AsthmaCritic needed an extension of the model to process the time-stamped data for the domain of asthma and COPD: event histories. In addition to an extension with event histories, AsthmaCritic also required an extension of the knowledge bases and data-processing functions, an implementation of specific functions to support integration in physicians working environment, and the means to offer the general practitioners control over the system's behavior. Although the prototype HyperCritic was not usable in daily practice, the generic model partially fit our needs and thus proved to be reusable.

Our design choices dealt with issues playing a role in acceptance as determined by CDSSs' degree of integration into physicians' working environment. We summarized these issues in a list that could be divided into two parts, each characterizing one of two identified user phases. First, 'Generating output', which deals with the hurdles a user has to overcome to get the system started (e.g., a user having to record additional data, or personally having to start the system). Second, 'Using output', which deals with the hurdles a user has to overcome to use the system's output (e.g., searching through heaps of information to obtain the desired support). The two user phases require different levels of user control. In the user phase 'Generating output'

user involvement should be *minimized* in order to optimize user acceptance – design choices should be made such that no control by the user is needed. In contrast, in the user phase 'Using output' user involvement should be *maximized* in order to optimize user acceptance – design choices should be made such that user control is enabled. Providing users with control when they use a system's output is one way to reconcile physicians' varying time and information needs. The summarized lists of issues may provide a handhold for system developers designing new critiquing systems aimed at integration in physicians' working environment, and researchers trying to gain insight into factors playing a role in system acceptance.

When making design choices, system developers are hampered by the lack of insight into the relationship between system characteristics and working environment characteristics. Researchers have attempted to characterize computerized decision-support systems (CDSSs), for example, by several dimensions or by specific information functions[3, 4]. Characterizing CDSSs along such dimensions helps to quickly identify a CDSS, but it does not help the clinician in knowing which system to choose to help him perform some task, and it does not help the system developer in optimizing design choices for system acceptance. Further research will be needed to, firstly, characterize working environments and, secondly, to perform studies that describe their relationship with system characteristics.

Guidelines, being summaries of large bodies of clinical evidence and having assimilated information such as policies, preferences, and resource availability, provide a good starting point for a CDSS's knowledge base. However, they have also shown to contain ambiguous and inconsistent information[5], making a knowledge engineer's task to translate a paper-based guideline into a formalized knowledge base error-prone. The dedication with which knowledge engineers perform their task directly determines the quality of the support offered by the resulting system, while quality control on the resulting knowledge base is limited. One way to reduce the vulnerability of the knowledge-acquisition process is to develop electronic guidelines that are to be used by CDSSs from the start. To support the development of electronic guidelines, researchers have started developing guideline implementation models. Ultimately, we feel that professional health-care organizations would have to organize the development of electronic guidelines, just as these organizations did with paper-based guidelines[6].

TESTING THE SYSTEM

Before any new tool can be put into practice, it needs to be tested to assess its quality and behavior. We knew from the simulation study that human reviewers were able to generate critiquing comments on routinely recorded data. We were now interested in the number and variety of critiquing comments that the system could generate. In addition, the system's robustness had to be tested to ensure the continuity of general practitioners' primary process.

In chapter 4 we describe how we assessed AsthmaCritic's ability to detect asthma and COPD patient records and generate patient-specific feedback in a laboratory setting. AsthmaCritic performed a retrospective analysis of routinely recorded data in over 100,000 electronic patient records from primary-care practices. We grouped generated feedback on contacts over one year into categories of comments by age group (<12 years and ≥12 years). AsthmaCritic detected 8.5% asthma and COPD patient records which is in line with results from Dutch registration networks (5-10%). During the study period, AsthmaCritic generated over 250,000 feedback comments (on average, 3.4 per patient contact) of 12 different categories. The study showed that the system, just like human reviewers, was able to select and critique records of patients with asthma or COPD and the system did not show unexpected functioning while working through these records. Therefore, we felt confident enough to pursue a study in daily practice.

The number of comments generated by AsthmaCritic in this study (on average, 3.4 per encounter) is considerable for daily practice. The acceptance of feedback, however, not only depends on the number of comments, but also on the kind of comments generated. The number and kind of comments vary dynamically, depending on recorded patient data and treatment decisions. On one hand, one can expect the number of generated comments to reduce if a physician decides to follow the guidelines (e.g., change his dosing schemas or start performing pulmonary function tests). On the other hand, the system may stimulate more complete recording of patient data, making the generation of more specific feedback possible. More specific comments could be perceived as being more useful, but they may be more likely to be wrong, thereby possibly jeopardizing the acceptance of all generated feedback. Further studies are needed to sort out the relationship between amount of

patient data, number of comments, feedback specificity, chances for false-positive feedback, and feedback acceptance.

Also, because the availability of patient data varies, a non-inquisitive CDSS should be prepared to deliver feedback on different levels of specificity, as is illustrated by the wide variety with which general practitioners use ICPC coding and the wide variety in specificity of feedback generated by AsthmaCritic. Figure 2 illustrates the variety with which general practitioners use their coding system to code a diagnosis asthma or COPD[7]. Physicians not used to recording much data should receive feedback at a somewhat generic level, but physicians recording many data should receive feedback taking that information into account.

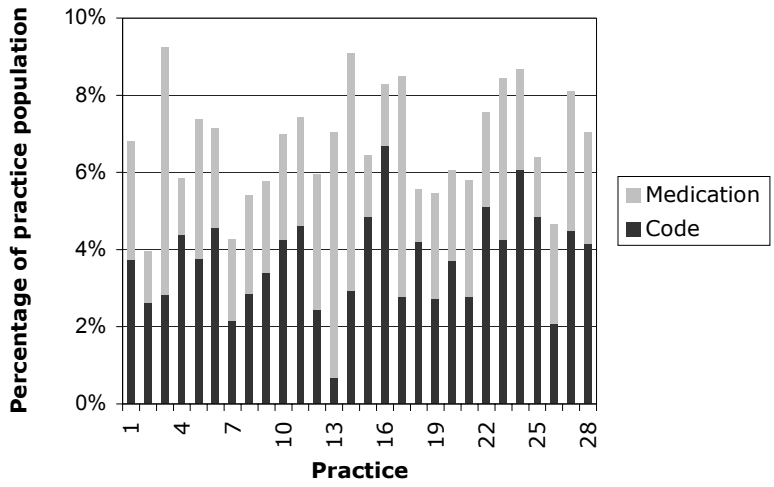


FIGURE 2.

CODING VARIABILITY. PERCENTAGE EXPLICITLY CODED OF THE ADULT PATIENT POPULATION TRIGGERED BY ASTHMACRITIC, PER PRACTICE.

A medically sound knowledge base is a prerequisite for a trustworthy CDSS. However, such a prerequisite does not guarantee feedback to be correct in all situations. Limitations on patient data availability and formalized medical knowledge create sources of uncertainty, which reduce feedback specificity and increase the risk for false-positiveness (a comment is generated while it should not have been generated). Therefore, physician interpretation on the clinical relevance of generated

comments will always be needed. Computers can only support human decision-making. Humans remain responsible for the decisions to be taken.

PUTTING ASTHMACRITIC INTO DAILY PRACTICE

The ultimate test for a CDSS is to be put into daily practice. In the busy routine of general practitioners, functionality determined by previously taken design choices has to prove its worth. To our knowledge, no-one has ever evaluated a non-inquisitive critiquing system by a randomised clinical trial in general practitioners' daily practice before.

To evaluate the effect of AsthmaCritic on general practitioners' behavior, we assessed the effect of the system on general practitioners' monitoring and treatment of asthma and COPD in daily practice. We conducted a randomised clinical trial with a five-month baseline and ditto intervention study period. We used the number of contacts, pulmonary function measurements, and five different kinds of prescriptions as our effect parameters. Our study showed that the system had been accepted and used by the general practitioners. Even though their hardware turned out to be older than average (causing an analysis to take, on average, about 30 seconds) they waited in about one-fifth of the cases for the feedback to be generated. The study also showed that the system changed the physicians' monitoring and, to a lesser extent, their treatment behaviour. In addition, the physicians changed their data recording habits in comparison with a control group, as could be seen by an increase in the ratio of measurements recorded in a structured fashion over measurements recorded in a combination of structured and unstructured (free-text) fashion.

AsthmaCritic generates feedback even when a patient comes for a different problem than asthma or COPD. This creates the opportunity to remind the physician, for example, to order a pulmonary function assessment irrespective of the reason for encounter. Generating feedback irrespective of the reason for encounter may improve asthma or COPD monitoring. However, if feedback is generated irrespective of the reason for encounter, the number of times the physician is exposed to comments increases, thereby possibly causing averse responses from the physician. Further research is needed to assess how system acceptance can be ensured while feedback exposure increases.

We believe that our design choices aimed at integrating AsthmaCritic into general practitioners' daily routine have been supportive in the system's effect on the physicians' behavior. What we do not know is which decisions precisely determined the acceptance of the system in the general practitioners' working environment. Our findings may stimulate other system developers to further explore the influence of specific choices on system acceptance. For example, the department of medical informatics at the Erasmus Medical Center Rotterdam is currently investigating the effect of trigger mode on physician acceptance of a cholesterol guideline implementation program.

STUDY LIMITATIONS

The first limitation of our study is that we evaluated the system in one region of The Netherlands – Delft-Westland. We, therefore, do not know if our findings will be applicable elsewhere.

We evaluated AsthmaCritic using one general practitioner information system only – ELIAS®. We, therefore, can not be sure that our findings also apply for non-inquisitive systems with other general practitioner information systems. AsthmaCritic has been implemented such that it would work independently of the kind of general practitioners information system. The system can do its job as long as the general practitioner information system is able to export patient data adhering to the current electronic patient data exchange standard MEDEUR[8]. However, to improve AsthmaCritic's performance we decided to bypass the time-consuming exchange interface that we built and limit ourselves to one information system.

We used a five-month baseline period and a five-month intervention period in the randomized clinical trial to assess AsthmaCritic's effect on physicians'' asthma and COPD management. Because of the five-month study period we could not study the system's possible effect on patient health or a possible fading-out of the observed effect on physicians' behavior. To study these two phenomena, a longer study period will be required.

Our choice for a non-inquisitive design limited ourselves in the choice of the level of abstraction of our effect parameters. We depended on the specificity of the available data in the electronic patient record, which does not necessarily match the specificity

of guideline recommendations. In other words, if data of some level of abstraction are needed to evaluate the effect of some intervention, it may be necessary to request those data during the intervention. Given our own design choice, we could not do so. Patient-specific indicators are needed to assess prescribing correctness independent of higher-level indicators in studies depending on electronic patient records [9]. Identifying patient-specific indicators is currently a hot topic for research.

FUTURE RESEARCH

At present, we do not know which information tools work best with what physician-working environment. We do not know which system characteristics determine a good fit between a system and the task and environment at hand. System developers, including ourselves, have to make their design choices based on a subjective evaluation of the intended system functionality in its future working environment. The lack of a theoretical model makes the scientific evaluation of CDSSs in relation to a working environment difficult. Further research is needed to be able to better understand and predict the relationship between system and environment[10].

Based on our own subjective evaluation, we made design choices aiming to optimize AsthmaCritic's chances to be accepted in general practice. Our field study showed that the system was used in daily practice. However, further research into the use of different components of the program and reasons why physicians reject or accept specific recommendations will help to gain insight into these decision-support systems' design issues.

Guidelines are sets of recommendations that guide the physician in treating his or her patients. Their recommendations will be appropriate in most cases. However, they can not predict and describe all circumstances. Therefore, the decision-making responsibility remains with the physician, who may decide to divert from the recommendations. The same is true for physicians using CDSSs based on such guidelines. However, critiquing a physician repeatedly that a patient is receiving twice the maximum dose, while the physician knows that the patient needs it, is very annoying. The acceptability of critiquing systems could be improved if physicians are able to store in an individualized knowledge base the fact that a specific patient needs a double dose. With an individualized knowledge base a physician can store patient-specific, personal, or local preferences regarding treatment decisions. The contra-argument is that creating an individualized knowledge base undermines the purpose

of a critiquing system – to point out to the physician that he or she diverts from established standards of medical behavior. We are very interested in studies that evaluate the use and acceptability of critiquing systems that incorporate the ability to adopt the system to personal preferences.

While individualized knowledge bases may help increase system acceptance by tailoring generated feedback, a possible expansion of the number of different clinical domains for which critiquing systems will be built, could lead to feedback overload. Given the varying time and information needs of general practitioners in daily practice, it is unlikely that a user will use all output generated by several domain-specific CDSSs during every patient encounter. Therefore, further research is needed in how to offer feedback in a manageable way in the busy routine of general practice.

CLOSING REMARKS

- General practitioners' electronic patient records contain sufficient patient data for human reviewers to critique general practitioners' monitoring and treatment of asthma and COPD.

- Because human reviewers do not update their own feedback when provided with additional patient data requested by themselves, a non-inquisitive design is the right choice for an integrated critiquing system.

- The analysis of issues underlying design choices for decison-support systems may lead to a better understanding of factors determining system acceptance.

- A non-inquisitive critiquing system is able to select electronic patient records of patients having asthma or COPD symptoms and subsequently critique general practitioners' monitoring and treatment.

- Because decision-support systems regard data with a limited scope, physician interpretation of comments will be needed to determine a critiquing system's clinical relevance.

- A non-inquisitive critiquing system changes physicians' monitoring and treatment of patients with asthma and COPD.

- The use of a non-inquisitive critiquing system changes general practitioners' recording behavior.

REFERENCES

1. van Ginneken AM, Verkoijen MJ. A multi-disciplinary approach to a user interface for structured data entry. Medinfo 2001;10(Pt 1):693-7.
2. Lei van der J, Musen MA. A model for critiquing based on automated medical records. Computers and Biomedical Research 1991;24:344-78.
3. Musen M, Shahar Y, Shortliffe EH. A structure for characterizing clinical decision-support systems. In: Shortliffe EH, Perreault LE, editors. Medical Informatics: Computer Applications in Health Care and Biomedicine. 2nd ed; 2000. p. 607-610.
4. Shiffman RN, Liaw Y, Brandt CA, Corb GJ. Computer-based guideline implementation systems: a systematic review of functionality and effectiveness. J Am Med Inform Assoc 1999;6(2):104-14.
5. van Wijk MA, Bohnen AM, van der Lei J. Analysis of the practice guidelines of the Dutch College of General Practitioners with respect to the use of blood tests. J Am Med Inform Assoc 1999;6(4):322-31.
6. NHG. 'NHG Standaarden' (Dutch College of General Practitioners Guidelines). Website: http://nhg.artsennet.nl/index.asp?s=1487;2002; Last accessed: 2002-06-20
7. Kuilboer MM, Mosseveld BMT, Lei van der J. AsthmaCritic: Het geintegreerd ondersteunen van het beleid voor patienten met astma/COPD op basis van het EMD. In: Huisarts, Specialist en het Elektronische Medisch Dossier.' Werken aan transmurele zorg'; 1997; Ede: Rotterdam: Vakgroep Medische Informatica; 1997. p. 49-59.
8. Vlug A. MEDEUR homepage. Website: http://www.eur.nl/fgg/mi/medeur/; Last accessed: 14 07 2002
9. Buetow SA, Sibbald B, Cantrill JA, Halliwell S. Prevalence of potentially inappropriate long term prescribing in general practice in the United Kingdom, 1980-95: systematic literature review [published erratum appears in BMJ 1997 Mar 1;314(7081):651]. British Medical Journal 1996;313(7069):1371-4.
10. Kaplan B. Evaluating informatics applications--some alternative approaches: theory, social interactionism, and call for methodological pluralism. Int J Med Inf 2001;64(1):39-56.

7

SAMENVATTING, DISCUSSIE EN TOEKOMSTIG ONDERZOEK

INLEIDING

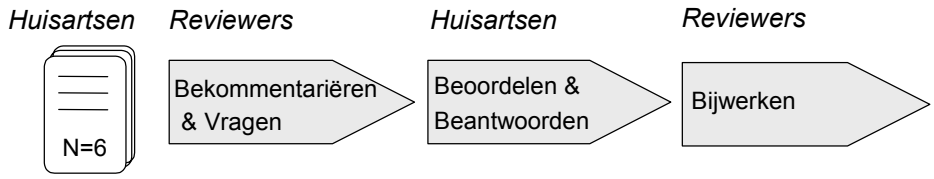
Dit onderzoek heeft als doel nader inzicht te verkrijgen in de potentie van kritieksystemen. Kritieksystemen zijn computerprogramma's die artsen kunnen helpen hun beroep volgens de huidige medische inzichten uit te oefenen. Deze systemen doen dat door het leveren van feedback op het handelen van huisartsen op basis van door hen in het elektronisch medisch dossier geregistreerde medische gegevens. Het onderzoek spitst zich toe op de haalbaarheid van een kritieksysteem en het effect ervan op het gedrag van huisartsen. Om de haalbaarheid en het effect van een kritieksysteem in de dagelijkse praktijk te onderzoeken hebben we voor het domein van astma en COPD (chronic obstructive pulmonary disease) zo'n systeem gesimuleerd, gebouwd, geïntegreerd met het elektronisch medisch dossier en geëvalueerd. De volgende alinea's geven een overzicht van elk van deze stappen en bevatten een reflectie op de verworven inzichten.

DE SIMULATIE VAN EEN KRITIEKSYSTEEM

Het verkrijgen van inzicht in de consequenties van ontwerpkeuzes door middel van het bouwen en evalueren van prototypen, is een tijdrovende inspanning. Als eerste stap hebben wij daarom een simulatiestudie uitgevoerd. Deze simulatiestudie stelde ons in staat om de belangrijkste kenmerken van het beoogde systeem te benoemen. Zo vereist een kritieksysteem elektronisch vastgelegde patiëntgegevens teneinde feedback te kunnen genereren. Ook al zijn de routinematig vastgelegde gegevens van huisartsen voldoende voor de dagelijkse praktijkvoering, dan betekent dat nog niet dat deze gegevens voldoende zijn om kritiek te kunnen leveren op hun handelen. Bovendien hebben artsen weinig tijd per consult. Onderbrekingen in de normale routine ten gevolge van verzoeken om extra informatie kunnen als storend worden ervaren, zelfs als deze onderbrekingen bedoeld zijn ter ondersteuning van het medisch handelen. Daarom wilden we eerst onderzoeken of de hoeveelheid in het elektronisch medisch dossier geregistreerde informatie voldoende was om kritiek te kunnen genereren. Als die informatie onvoldoende zou blijken, wilden we weten welke informatie ontbrak en hoe belangrijk deze ontbrekende informatie was voor het genereren van opmerkingen.

In de simulatiestudie, zoals beschreven in Hoofdstuk 2, vroegen we vier 'reviewers' (twee huisartsen en twee specialisten, allen gespecialiseerd in astma en COPD) om de rol van computer te spelen en opmerkingen te genereren op basis van de

elektronische medische dossiers van patiënten met astma of COPD. We vroegen drie huisartsen om de rol van gebruiker te spelen, de opmerkingen te beoordelen en eventuele ontbrekende informatie (zoals aangegeven door de reviewers) te verstrekken. Tenslotte vroegen we de vier reviewers om hun opmerkingen zo nodig te herzien op basis van de alsnog geleverde ontbrekende informatie. Figuur 1 laat de opeenvolgende stappen van de simulatiestudie zien.



FIGUUR 1.

DE DRIE OPEENVOLGENDE STAPPEN IN DE SIMULATIESTUDIE. VIER REVIEWERS ANALYSEERDEN ZES MEDISCHE DOSSIERS (MET IN TOTAAL 87 CONTACTEN). DE REVIEWERS LEVERDEN KOMMENTAAR EN VROEGEN OM NADERE INFORMATIE ALS ZE DAT NODIG ACHTTEN ('ONTBREKENDE INFORMATIE'). DE HUISARTSEN BEOORDEELDEN DE OPMERKINGEN EN VERSTREKTEN DE ONTBREKENDE INFORMATIE. ALS DE INFORMATIE NIET BESCHIKBAAR WAS WERD HEN GEVRAAGD OM AAN TE GEVEN WAAROM NIET. TENSLOTTE BEKEKEN DE REVIEWERS DE OPMERKINGEN OPNIEUW EN BRACHTEN EVENTUEEL VERBETERINGEN AAN OP BASIS VAN DE ALSNOG BESCHIKBAAR GEKOMEN INFORMATIE.

De studie liet zien dat er verschillende soorten opmerkingen gegenereerd konden worden en dat veel van de ontbrekende informatie toch beschikbaar bleek nadat daar specifiek om gevraagd was. De reviewers lieten echter driekwart van hun opmerkingen onveranderd nadat de ontbrekende informatie beschikbaar kwam. Op basis hiervan besloten we een "non-inquisitive" systeem te ontwerpen. "Non-inquisitive" betekent dat het systeem geen extra informatie van de arts vraagt buiten de gegevens die de arts reeds routinematig registreert. Samengevat konden we zeggen:

HET GEBRUIK VAN HET ELEKTRONISCH MEDISCH DOSSIER ALS ENIGE GEGEVENSBRON VOOR EEN KRITIEKSYSTEEM VOOR ASTMA EN COPD IS HAALBAAR, WAARDOOR EEN "NON-INQUISITIVE" SYSTEEMONTWERP MOGELIJK IS.

Tijdens de simulatiestudie vroegen we aan de huisartsen waarom de (volgens de reviewers) ontbrekende informatie niet geregistreerd was. In de meeste gevallen

bleek de informatie niet expliciet geregistreerd dan wel elders in het dossier opgenomen te zijn (niet noodzakelijkerwijs toegankelijk voor een beslissingsondersteunend computer programma). Deze redenen voor het ontbreken van informatie illustreren het spanningsveld tussen de gegevensbehoefte van de huisartsen en van de beoordelaars (critici). Om artsen te stimuleren meer expliciet en gestructureerd te registreren zou de wijze van gestructureerd invoeren van gegevens vereenvoudigd moeten worden. Dit is echter lastig vanwege het daarbij optredende spanningsveld tussen complexiteit en helderheid en gebruiksgemak[1].

HET ONTWERP VAN EEN GEÏNTEGREERD SYSTEEM

In Hoofdstuk 3 zijn we nader ingegaan op de ontwerpkeuzes die we, naast de keuze voor het “non-inquisitive” systeem, voor AsthmaCritic gemaakt hebben. Tevens gaan we in op diverse aspecten die een rol spelen in de acceptatie van het systeem. Ons doel was een systeem te ontwerpen dat geïntegreerd is met de dagelijkse praktijkvoering van de huisarts. Met ‘geïntegreerd’ bedoelen we enerzijds een systeem dat zijn gegevens ontvangt van het huisartseninformatiesysteem en anderzijds een systeem dat past in de dagelijkse praktijkvoering van huisartsen.

Voor de implementatie van AsthmaCritic hebben we het algemene kritiekmodel van Van der Lei uitgebreid om aan de eisen van de dagelijkse praktijk te kunnen voldoen. Het algemene kritiekmodel ondersteunt de integratie met een elektronisch medisch dossier op dataniveau. Hiertoe maakt het gebruik van gebeurtenissen (“events”) die de brug slaan tussen het elektronisch medisch dossier en het kritieksysteem. Van der Lei heeft, om zijn model te testen, een prototype gebouwd: HyperCritic. HyperCritic behelst het domein van hypertensie en is alleen getest in een laboratoriumsituatie. Het domein van de aandoeningen astma en COPD is complexer dan het domein van de risicofactor hypertensie. Om bijvoorbeeld de ernst van de achteruitgang van een astma of COPD-patiënt vast te kunnen stellen, moet het systeem een sequentie van symptomen, voorschriften en longfunctiemetingen in hun onderlinge samenhang kunnen beoordelen. Voor AsthmaCritic was het daarom noodzakelijk om het generieke model uit te breiden met een mogelijkheid om tijdreeksen te verwerken: de zogenaamde “event-histories”. Daarnaast was voor AsthmaCritic een uitbreiding nodig van de kennisbestanden en de programmatuur. Bovendien waren specifieke functies nodig om de integratie in de dagelijkse praktijk te ondersteunen en de mogelijkheid voor artsen om de controle over het systeem te behouden. Alhoewel het

prototype HyperCritic niet bruikbaar was in de dagelijkse praktijk, bleek het generieke model gedeeltelijk herbruikbaar.

De mate waarin een beslissingsondersteunend systeem geïntegreerd wordt in de dagelijkse praktijk bepaalt in grote mate de acceptatie van een dergelijk systeem. Onze ontwerpkeuzes hadden betrekking op aspecten die een rol spelen bij deze acceptatie. Samenvattend zijn deze aspecten in te delen in twee hoofdcategorieën, die elk één gebruikersfase karakteriseren. De eerste, 'het genereren van output', heeft betrekking op de hindernissen die een gebruiker moet overwinnen om het systeem op te starten. Het aspect 'data-invoer' speelt bijvoorbeeld een rol bij de keus of er gegevens separaat ingevoerd zouden moeten worden – het apart in moeten voeren van gegevens kan een extra hindernis opwerpen voor de potentiële gebruiker bij de acceptatie van het systeem. De tweede gebruikersfase, 'het gebruiken van output', heeft betrekking op de hindernissen die de gebruiker moet overwinnen om de output van het systeem te kunnen gebruiken. Zo kan het moeten doorzoeken van veel informatie voordat de beoogde ondersteuning verkregen wordt een hindernis zijn die te maken heeft met de aspecten 'informatie differentiatie' en met 'informatie doelmatigheid'. De twee gebruikersfasen vereisen verschillende niveaus van controle. In de gebruikersfase 'het genereren van output' is het belangrijk dat van de gebruiker een *minimale* inspanning gevraagd wordt om tot een optimale acceptatie te komen. In de gebruikersfase 'het gebruiken van output' daarentegen, moet de controle van de gebruiker *maximaal* zijn om de acceptatie van het systeem te bevorderen. Gebruikers de mogelijkheid geven om de output van het systeem aan de eigen wensen aan te passen is een methode om tegemoet te komen aan de in de huisartsenpraktijk wisselende informatiebehoefte en variërende tijdsdruk. Onze lijst met aspecten kan systeemontwerpers ondersteunen bij het ontwerpen van nieuwe kritieksystemen voor de dagelijkse praktijk. Daarnaast kan het onderzoekers ook helpen inzicht te verkrijgen in factoren die een rol spelen bij de acceptatie van een systeem.

Ontwerpers van systemen worden bij het maken van ontwerpkeuzes beperkt door het ontbreken van inzicht in de relatie tussen systeemkenmerken en de kenmerken van de werkomgeving. Onderzoekers hebben geprobeerd om beslissingsondersteunende systemen te karakteriseren aan de hand van verschillende dimensies of verschillende informatie-functies[3, 4]. Het karakteriseren van een beslissingsondersteunend systeem aan de hand van zulke dimensies helpt de ontwerper echter niet bij het optimaliseren van ontwerpkeuzes met betrekking tot systeem acceptatie. Ook helpt het de clinicus

niet bij het kiezen van een systeem dat het beste aansluit bij de uit te voeren taken. Nader onderzoek is nodig, enerzijds om werkomgevingen te karakteriseren en anderzijds om inzicht te krijgen in de relatie tussen de kenmerken van de werkomgeving en de kenmerken van het systeem.

Richtlijnen vormen een goed startpunt om het kennisbestand van een beslissingsondersteunend systeem mee op te bouwen. Een richtlijn is een samenvatting van grote hoeveelheden “clinical evidence” en integreert bovendien informatie over beleid, voorkeuren en beschikbaarheid van informatiebronnen. Echter, richtlijnen blijken ook ambivalente en inconsistente informatie te bevatten[5], waardoor de “knowledge engineer” bij het vertalen van de papieren richtlijnen naar een geformaliseerd bestand fouten kan maken (een geformaliseerd bestand is een bestand dat geschikt gemaakt is voor computerverwerking). De toewijding waarmee “knowledge engineers” hun taak uitvoeren bepaalt de kwaliteit van de ondersteuning die door het systeem gegeven wordt. Er is echter maar een beperkte kwaliteitscontrole op het resulterende kennisbestand mogelijk. Eén manier om de kwetsbaarheid van het kennisverwervingproces te beperken is om elektronische richtlijnen te ontwikkelen die rechtstreeks gebruikt kunnen worden in een beslissingsondersteunend systeem. Om de ontwikkeling van elektronische richtlijnen te bevorderen zijn onderzoekers begonnen met het ontwikkelen van richtlijn implementatiemodellen. Wij zijn van mening dat uiteindelijk de professionele gezondheidszorgorganisaties de taak op zich moeten nemen om deze elektronische richtlijnen te ontwikkelen, net zoals ze dat gedaan hebben met de papieren richtlijnen[6].

HET TESTEN VAN HET SYSTEEM

Voordat een nieuw hulpmiddel in de praktijk gebruikt kan worden, moet het op kwaliteit en gedrag getest worden. Wij wisten op basis van de simulatiestudie dat menselijke reviewers in staat waren om kritiek te leveren op basis van routinematig vastgelegde gegevens. Ons volgende onderzoeksdoel was het vaststellen van het aantal en de soort opmerkingen dat het systeem kon genereren. Daarnaast moest de robuustheid van het systeem getest worden om de continuïteit van het primaire proces in de huisartsenpraktijk te kunnen garanderen.

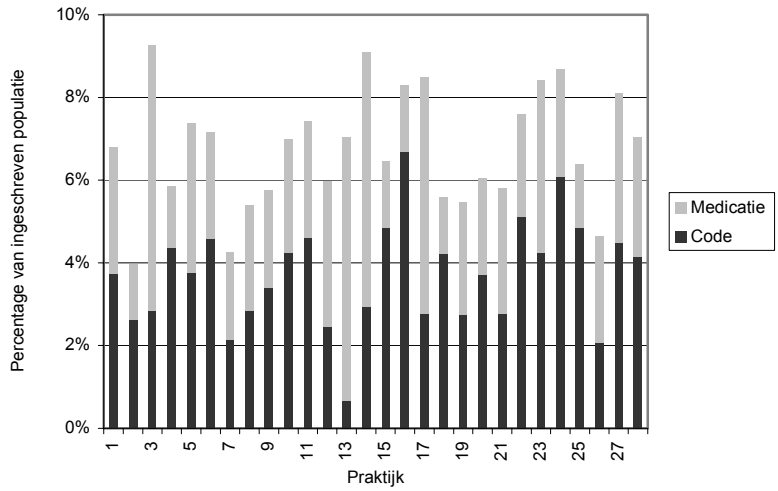
In hoofdstuk 4 stelden we in een laboratoriumsetting het vermogen van AsthmaCritic vast om astma en COPD dossiers te selecteren en om patiëntspecifieke feedback te

genereren. We lieten AsthmaCritic een retrospectieve analyse uitvoeren van routinematig vastgelegde gegevens in ruim 100.000 elektronische medische dossiers uit diverse huisartsenpraktijken. AsthmaCritic becommentarieerde de contacten over een periode van een jaar. Wij groepeerden de feedback op basis van leeftijd (jonger dan 12 jaar en 12 jaar en ouder). AsthmaCritic vond 8,5% astma- en COPD-dossiers, hetgeen overeenkomt met het prevalentiecijfer uit de Nederlandse registratienetwerken (5-10%). AsthmaCritic genereerde gedurende de studieperiode meer dan 250.000 opmerkingen (gemiddeld 3,4 per consult) verdeeld over 12 verschillende categorieën. Het onderzoek liet zien dat het systeem, evenals de menselijke reviewers, in staat was om dossiers van patiënten met astma- of COPD-symptomen te selecteren en te becommentariëren. Het systeem liet geen onverwachte functionaliteit zien tijdens de bewerking van deze dossiers. Op basis hiervan zagen we een veldstudie met AsthmaCritic in de dagelijkse huisartsenpraktijk met vertrouwen tegemoet.

Het aantal opmerkingen dat AsthmaCritic in deze studie genereerde (gemiddeld 3,4 per contact) is aanzienlijk voor de dagelijkse praktijk. De acceptatie van feedback hangt echter niet alleen af van het aantal opmerkingen, maar ook van het soort opmerkingen dat gegenereerd wordt. Het aantal en soort opmerkingen varieert dynamisch, afhankelijk van de hoeveelheid geregistreerde patiëntgegevens en de behandelkeuzes. Aan de ene kant kan men verwachten dat het aantal opmerkingen afneemt als een arts besluit om de richtlijnen te volgen (b.v. een verandering van het toegepaste doseringsschema of het laten uitvoeren van een longfunctietest). Aan de andere kant kan het systeem ertoe aanzetten dat er meer gestructureerd en gecodeerd geregistreerd wordt, hetgeen het aanmaken van meer specifieke feedback mogelijk maakt. Nader onderzoek is nodig om de relatie tussen de hoeveelheid gestructureerd geregistreerde patiëntgegevens, het aantal opmerkingen, de specificiteit van de feedback, de kans op het genereren van foutpositieve feedback en de mate van acceptatie van de feedback vast te stellen.

Omdat de beschikbaarheid van patiëntgegevens varieert, moet een “non-inquisitive” beslissingsondersteunend systeem er op voorbereid zijn om feedback te geven op verschillende niveaus van specificiteit. De noodzaak hiervan wordt geïllustreerd door de grote variatie in de mate waarin huisartsen ICPC-codes gebruiken en de grote variatie in de specificiteit van de opmerkingen die door AsthmaCritic gegenereerd worden. Figuur 2 laat de variatie zien in de mate waarin huisartsen hun

coderingssysteem gebruiken om de diagnose astma of COPD te registreren[7]. Artsen die niet gewend zijn om veel gegevens gecodeerd te registreren moeten feedback krijgen op een algemeen niveau, terwijl artsen die wel veel gegevens gecodeerd registreren feedback moeten krijgen op een specifiek niveau.



FIGUUR 2.

VARIATIE IN CODERINGSGRAAD. HET PERCENTAGE EXPLICIETE CODERINGEN PER PRAKTIJK VOOR VOLWASSEN PATIËNTEN DIE DOOR ASTHMACRITIC ALS ASTMA OF COPD-PATIËNT GEKENMERKT ZIJN.

Een medisch solide kennisbestand is een voorwaarde voor een vertrouwenwekkend beslissingsondersteunend systeem. Zo'n voorwaarde garandeert echter niet dat de feedback in alle situaties juist is. Beperkingen in de beschikbaarheid van patiëntgegevens en geformaliseerde medische kennis creëren bronnen van onzekerheid, die de specificiteit van de feedback verminderen en het risico van foutpositieve feedback doen toenemen. Een foutpositieve opmerking is een opmerking die ten onrechte gemaakt wordt. Dientengevolge blijft het noodzakelijk dat de arts de klinische relevantie van de opmerkingen interpreteert. Computers kunnen de menselijke besluitvorming alleen ondersteunen. Mensen blijven verantwoordelijk voor de te nemen beslissingen.

HET IN GEBRUIK NEMEN VAN ASTHMACRITIC IN DE DAGELIJKSE PRAKTIJK

De ultieme test voor de haalbaarheid van een beslissingsondersteunend systeem is het in gebruik nemen in de dagelijkse praktijk. In de drukke routine van de huisartsenpraktijk moet de functionaliteit die het resultaat is van eerder genomen ontwerpkeuzes zijn waarde bewijzen. Voor zover wij weten zijn er nog niet eerder “non-inquisitive” kritieksystemen getest door middel van een gerandomiseerd klinisch onderzoek in de huisartsenpraktijk.

Om het effect van AsthmaCritic op het monitoren en behandelen van astma en COPD-patiënten door huisartsen in de dagelijkse praktijk te bepalen, werd na stratificatie naar solo- en groepspraktijken een gerandomiseerd onderzoek uitgevoerd. Het onderzoek vond plaats in 32 praktijken in de regio Delft, waarin 40 huisartsen praktiseerden. De onderzoeksperiode bestond uit een nulmeting van vijf maanden en een even lange studieperiode. We gebruikten het aantal contacten, de longfunctiemetingen en vijf verschillende soorten voorschriften als onze effectparameters. Onze studie liet zien dat het systeem geaccepteerd en gebruikt werd door de huisartsen. Hoewel de hardware van de huisartsen gemiddeld genomen vrij oud bleek (waardoor een analyse gemiddeld 30 seconden duurde), wachtten de huisartsen in éénvijfde van de gevallen tot de feedback gegenereerd was. Het onderzoek liet ook zien dat het systeem veranderingen teweegbracht in het monitoren van de patiënt en, in mindere mate, in de behandeling. Bovendien veranderden de artsen hun registratiegewoontes ten opzichte van de controlegroep.

AsthmaCritic genereert ook feedback als patiënten voor andere problemen dan astma of COPD komen. Dit biedt bijvoorbeeld de mogelijkheid om de arts te herinneren aan het bepalen van de longfunctie, ongeacht de reden van het contact. Het genereren van feedback onafhankelijk van de reden voor het contact, kan het monitoren van astma en COPD-patiënten ondersteunen. Echter het genereren van opmerkingen onafhankelijk van de reden van het contact resulteert ook in een stijging van het aantal momenten dat een arts geconfronteerd wordt met opmerkingen. Nader onderzoek is nodig om vast te stellen hoe de acceptatie van het systeem gewaarborgd kan worden bij een toename van de hoeveelheid feedback.

We denken dat onze ontwerpkeuzes, die gericht waren op het integreren van AsthmaCritic in de dagelijkse huisartsenpraktijk, belangrijk waren voor het geobserveerde effect op het gedrag van de artsen. Wat we niet weten is welke

keuzes precies de acceptatie van het systeem hebben bevorderd. Nader onderzoek zal moeten uitwijzen welk effect specifieke ontwerpkeuzes hebben op de acceptatie van beslissingsondersteunende systemen. .

BEPERKINGEN VAN DE STUDIE

De eerste beperking van ons onderzoek is dat we het systeem testten in één Nederlandse regio: Delft-Westland. Op grond daarvan weten we niet of onze bevindingen ook elders toepasbaar zijn.

Wij onderzochten de werking van AsthmaCritic bij slechts één huisartsinformatiesysteem (HIS): ELIAS®. Daarom kunnen we er niet zeker van zijn dat onze resultaten ook toepasbaar zijn voor 'non-inquisitive' systemen geïntegreerd in andere HISsen. AsthmaCritic is zo ontworpen dat het onafhankelijk van het type HIS kan functioneren. Het systeem kan z'n werk doen als het HIS in staat is om patiëntgegevens te exporteren volgens de Nederlandse standaard voor uitwisseling van elektronische patiëntgegevens MEDEUR[8]. Echter, om de performance van AsthmaCritic te verbeteren, hadden we besloten om het tijdrovende proces van het bouwen van uitwisselingsinterfaces over te slaan en ons te beperken tot slechts één HIS.

Voor de gerandomiseerde klinische studie gebruikten we een nulmeting van vijf maanden en een even lange interventieperiode om het effect van AsthmaCritic te testen op de manier waarop huisartsen astma en COPD-patiënten monitoren en behandelen. Als gevolg van deze relatief korte periodes konden we een potentieel effect op de gezondheid van de patiënt niet testen en evenmin vaststellen of er een 'fading-out' van het effect op het gedrag van de huisarts zou ontstaan. Om deze twee fenomenen te bestuderen is een langere studieperiode noodzakelijk.

Door onze keuze voor een "non-inquisitive" ontwerp beperkten we onszelf in de keuze van het abstractieniveau van de effectparameters. We waren afhankelijk van de specificiteit van de beschikbare gegevens in het elektronische patiëntendossier. Deze specificiteit komt niet noodzakelijkerwijs overeen met de specificiteit zoals omschreven in de aanbevelingen van de richtlijnen. Met andere woorden, als gegevens van een bepaald abstractieniveau noodzakelijk zijn om het effect van de interventie te evalueren, kan het noodzakelijk zijn om deze gegevens gedurende de interventie te laten registreren. Gezien onze eigen ontwerpkeuze was dat voor ons

niet mogelijk. Om de correctheid van het voorschrijfgedrag vast te stellen zijn, onafhankelijk van indicatoren op hoger niveau, patiëntspecifieke indicatoren nodig[9]. Het vaststellen van patiëntspecifieke indicatoren staat momenteel hoog op de onderzoeksagenda.

TOEKOMSTIG ONDERZOEK

Voor verschillende werkomgevingen zijn verschillende 'informatiegereedschappen' nodig. Op dit moment is niet bekend welke 'informatiegereedschappen' het beste passen bij welk type werkomgeving. We weten niet welke systeemkarakteristieken een goede match zullen creëren tussen een systeem en de uit te voeren taak in zijn specifieke omgeving. Systeemontwikkelaars, waartoe wijzelf ook behoren, baseren hun keuze op een subjectieve evaluatie van de beoogde systeemfunctionaliteit in haar toekomstige werkomgeving. Het ontbreken van een theoretisch model maakt de wetenschappelijke evaluatie van een beslissingsondersteunend systeem in relatie tot een werkomgeving moeizaam. Nader onderzoek is nodig om de relatie tussen systeem en omgeving beter te kunnen begrijpen en voorspellen[10].

Op basis van onze eigen subjectieve evaluatie maakten we keuzes die gericht waren op het optimaliseren van de acceptatie van AsthmaCritic in de huisartsenpraktijk. Onze veldstudie liet zien dat het systeem bruikbaar was in de dagelijkse praktijk. Nader onderzoek naar het gebruik van verschillende componenten van het programma en naar de redenen waarom artsen sommige aanbevelingen verwerpen of accepteren kan meer inzicht geven in aspecten van ontwerpkeuzes die een rol spelen bij de acceptatie van beslissingsondersteunende systemen.

Richtlijnen zijn sets van aanbevelingen die artsen richting geven bij de behandeling van hun patiënten. Hoewel de aanbevelingen in de meeste gevallen juist zullen zijn kunnen ze niet in alle omstandigheden voorzien. Daarom blijft de verantwoordelijkheid voor de besluitvorming bij de arts liggen, die kan besluiten om in voorkomende gevallen af te wijken van de richtlijn. Echter, het kan erg storend zijn als een arts bijvoorbeeld meerdere malen de opmerking krijgt dat een patiënt twee maal de toegestane dosis krijgt, terwijl de arts weet dat het een voor die patiënt adequate dosering is. De mogelijkheid dat een arts een individueel kennisbestand zou kunnen opbouwen met patiëntspecifieke, persoonlijke of locale voorkeuren ten aanzien van de behandeling zou de acceptatie van een kritieksysteem mogelijk kunnen bevorderen. Daartegen pleit dat het maken van een individueel kennisbestand het

doel van een kritieksysteem ondermijnt: de arts moet er juist op gewezen worden dat hij of zij afwijkt van de richtlijn. Nader onderzoek moet uitwijzen of en hoe een individueel kennisbestand gebruikt zal worden en of deze voorziening ertoe kan bijdragen dat kritieksystemen beter geaccepteerd zullen worden.

Uitbreiding van het aantal verschillende domeinen waarvoor een kritieksysteem gebruikt wordt kan ertoe leiden dat er een “feedback-overload” ontstaat. Gezien de variatie in beschikbare tijd en behoefte aan informatie van de huisartsen in de dagelijkse praktijk, is het onwaarschijnlijk dat een gebruiker tijdens een consult alle output zal gebruiken die door diverse domeinspecifieke beslissingsondersteunende systemen aangemaakt wordt. Daarom is nader onderzoek nodig naar acceptabele manieren om grotere hoeveelheden feedback aan te bieden in de drukke routine van de huisartsenpraktijk.

SAMENVATTENDE SLOTOPMERKINGEN

- Elektronische medische dossiers van patiënten van huisartsen bevatten voldoende informatie voor menselijke reviewers om kritiek te leveren op het monitoren en behandelen van astma en COPD-patiënten door huisartsen.

- Omdat menselijke reviewers het niet nodig achtten om hun feedback te veranderen na het beschikbaar komen van aanvullende informatie waar ze zelf om gevraagd hadden, is een “non-inquisitive” systeem een verantwoorde keuze bij het ontwerpen van een geïntegreerd kritieksysteem.

- De analyse van de achterliggende aspecten van ontwerpkeuzes die gemaakt moeten worden bij de bouw van beslissingsondersteunende systemen, draagt bij tot een betere onderbouwing van deze ontwerpkeuzes. Dit verhoogt de kans op acceptatie van zo’n systeem.

- Een “non-inquisitive” kritieksysteem is in staat om elektronische dossiers van patiënten met astma of COPD te selecteren en vervolgens opmerkingen te genereren met betrekking tot de handelswijze van de huisarts.

- Omdat een beslissingsondersteunend systeem slechts een beperkt zicht heeft op de informatie die potentieel relevant is voor de behandeling van een patiënt, blijft het noodzakelijk dat een arts de klinische relevantie van de feedback beoordeelt.

- Een “non-inquisitive” kritieksysteem verandert de wijze waarop artsen astma en COPD-patiënten monitoren en behandelen.

- Het gebruik van een “non-inquisitive” kritieksysteem verandert het registratiegedrag van huisartsen.

REFERENTIES

1. van Ginneken AM, Verkoijen MJ. A multi-disciplinary approach to a user interface for structured data entry. Medinfo 2001;10(Pt 1):693-7.
2. Lei van der J, Musen MA. A model for critiquing based on automated medical records. Computers and Biomedical Research 1991;24:344-78.
3. Musen M, Shahar Y, Shortliffe EH. A structure for characterizing clinical decision-support systems. In: Shortliffe EH, Perreault LE, editors. Medical Informatics: Computer Applications in Health Care and Biomedicine. 2nd ed; 2000. p. 607-610.
4. Shiffman RN, Liaw Y, Brandt CA, Corb GJ. Computer-based guideline implementation systems: a systematic review of functionality and effectiveness. J Am Med Inform Assoc 1999;6(2):104-14.
5. van Wijk MA, Bohnen AM, van der Lei J. Analysis of the practice guidelines of the Dutch College of General Practitioners with respect to the use of blood tests. J Am Med Inform Assoc 1999;6(4):322-31.
6. NHG. 'NHG Standaarden' (Dutch College of General Practitioners Guidelines). Website: http://nhg.artsennet.nl/index.asp?s=1487;2002; Last accessed: 2002-06-20
7. Kuilboer MM, Mosseveld BMT, Lei van der J. AsthmaCritic: Het geintegreerd ondersteunen van het beleid voor patiënten met astma/COPD op basis van het EMD. In: Huisarts, Specialist en het Elektronische Medisch Dossier.' Werken aan transmurele zorg'; 1997; Ede: Rotterdam: Vakgroep Medische Informatica; 1997. p. 49-59.
8. Vlug A. MEDEUR homepage. Website: http://www.eur.nl/fgg/mi/medeur/; Last accessed: 14 07 2002
9. Buetow SA, Sibbald B, Cantrill JA, Halliwell S. Prevalence of potentially inappropriate long term prescribing in general practice in the United Kingdom, 1980-95: systematic literature review [published erratum appears in BMJ 1997 Mar 1;314(7081):651]. British Medical Journal 1996;313(7069):1371-4.
10. Kaplan B. Evaluating informatics applications--some alternative approaches: theory, social interactionism, and call for methodological pluralism. Int J Med Inf 2001;64(1):39-56.

DANKWOORD

Iemand beweert met de regelmaat van de klok dat het schrijven van een proefschrift “een exercitie in eenzaamheid is’’. Gezien de lijst van mensen die meegeholpen hebben aan de totstandkoming van dit boekje zou dat onmogelijk moeten zijn! In dit hoofdstuk wil ik aldus iedereen die bijgedragen heeft, zij het financieel, inhoudelijk, tekstueel, fysiek of emotioneel, bedanken voor zijn of haar hulp. Mijn dank voor jullie is ontzettend groot!

Allereerst wil ik het Nederlands Astma Fonds hartelijk danken voor de financiële middelen die zij ter beschikking hebben gesteld voor dit experimentele werk. Ook dit soort onderzoek draagt bij tot verder inzicht in en uiteindelijk verbetering van de zorg voor patiënten met astma en COPD.

Johan, jij stond in feite aan de wieg van mijn medische informatica loopbaan, jij stimuleerde me om de gang naar Stanford vooral wél te maken. Daarna togen we aan het AsthmaCritic-werk. We hebben beiden niet gedacht dat het zo’n klus zou worden. Gelukkig is het een geslaagde klus geworden. En ook al maakte jij je wel eens zorgen, je hebt er achter de schermen altijd voor gezorgd dat ik door kon gaan. En tsja, ik zal nooit zo ‘bondig’ worden als jij, maar wat ik tegenwoordig produceer is maar half zo weids als toen we ooit begonnen. Bedankt!

Marc, van collega’s werden we partners en werden we begeleider-promovendus. Dankzij jou heeft veel kunnen plaatsvinden wat anders nooit op deze manier gelukt was. Nu ruim anderhalf jaar geleden zorgde jij voor een nieuwe continuïteit in m’n werk. Ik ben je ontzettend dankbaar voor je tomeloze inzet en je onwrikbare vertrouwen in de goede afloop. We hebben samen veel plezier gehad en ik vond de, soms heftige, discussies heerlijk. Weet verder nog dat je turn-over tijd voor het becommentariëren van nieuwe versies onovertroffen is. En mocht je je gaan vervelen, ik heb nog wel wat materiaal op de plank liggen!

Mees, collega, compagnon, steunpilaar, maatje van de afgelopen jaren. Zoals ik al eens zei; volgens mij moeten er twee namen op de kaft staan. Er is geen beginnen aan om de energie en de uren - by day or by night – te gaan tellen, maar we hebben er wel iets moois van gemaakt! Mees, ontzettend bedankt voor alles wat je voor me gedaan hebt. Het is onbeschrijfbaar.

Emiel, vanwege de vertragingen sta jij niet meer in mijn boekje als promotor. Ik ben heel erg blij dat je wel bij de promotie zult zijn. Ook jij vormde een rode draad door de afgelopen jaren heen. Een ervaren draad met een kritische blik op ‘die dokters’ en de maatschappij. Dat moest veel beter kunnen. En misschien kon AsthmaCritic daar wel bij helpen. En kon jij mij helpen AsthmaCritic te maken. Dankjewel voor al je steun!

Johan de Jongste, Ben Ponsioen, Emiel van der Does, Shelley Overbeek
Jullie vormden het ‘medical content board’, het klinisch geweten van AsthmaCritic. Dank voor jullie moed en energie om je door alle versies van het kennisbestand heen te werken en de volgende versie van een artikel te bekijken. Het is een fijne samenwerking geweest.

Jifke, Anke, en Martine, kamergenoten en vriendinnen, dames in de medische informatica. Heerlijke discussies, het delen van de ups en downs van het AIO-leven, het leven met of juist zonder kinderen, wat te doen na de promotie en hoe de goede balans te vinden. Jullie stonden altijd klaar om bij te springen – ontzettend bedankt! Ik hoop dat we het nog lang zullen kunnen volhouden.

Jan van Bemmel, eerst als promotor, nu als rector bij mijn promotie. De afgelopen jaren klonk regelmatig de vraag: “en, wanneer ga je promoveren?”. Jan, hou je vast, het gaat er echt van komen. Ik ben blij dat ik deel heb mogen uitmaken van ‘jouw vakgroep’!

Jan Grashuis, eens een mentor, altijd een mentor! Dank voor je steun, zowel in diptijden als in Apple-tijden. Ik heb zowaar een verse i-book gesignaleerd op de vakgroep, dus wie weet raakt de techniek niet geheel verloren. Ik hoop dat Arabesk gouden tijden mag beleven, dan kan ik nog vaak komen ‘spelen’.

Astrid, orca’s, badminton avondjes, en bridgen. Fijne herinneringen aan speciale momenten. Helaas was het zo dat hoe groter mijn gezin werd, hoe minder tijd wij voor elkaar hadden. En inderdaad, na die eerste keer Orca Survey heeft een tweede keer er nog niet ingezeten. Maar wat niet is, kan nog komen. Zelfs een boek over het leven van een Orca als je tot dik over je oren in het werk zit!

Albert, wat is er sterker dan de combinatie techniek en filosofie? Ik zou het niet weten, maar ik weet wel dat een andere blik op ons vak tot heel verfrissende

inzichten kan leiden. Ik heb genoten van de (kamerbrede) discussies die zo nu en dan mogelijk waren, ik heb bewondering voor je tomeloze energie en je vriendelijke geduld met Jan en Alleman, en ik ben ontzettend benieuwd wat jij nog allemaal gaat uitspoken. Dank voor je vriendschap en hou me op de hoogte!

Peter Moorman, Desiree, Roos, Joop, Katia, en alle andere vakgroepsgenoten – we vormen met z'n allen een goed team. Bedankt voor alle gezelligheid, hulp en vriendschap. Zonder goede collega's is goed werk blijven leveren onmogelijk!

Thom Enneking, promoveer zelf nou eens, dan had ik je in mijn commissie mogen hebben. Dank voor het testwerk dat je aan AsthmaCritic gedaan hebt en voor de fantastische zomer in je praktijk waarin ik weer eens even aan den lijve mocht ondervinden wat een geweldig vak jij hebt.

Ook buiten de vakgroep en AsthmaCritic om wil ik een paar mensen speciaal bedanken die direct of indirect voorwaardenscheppend zijn geweest bij de totstandkoming van dit boekje.
Jan-Maarten Wit, jij staat feitelijk aan de basis van mijn wetenschappelijke carrière, je was mijn mentor van het eerste uur en hebt menig uurtje geholpen bij de voorbereidingen van de gang naar Stanford. Ik ben je veel schuldig! Ik wil jou en Birgit enorm bedanken voor al jullie steun en vriendschap. Ik hoop dat de toekomst nog veel meer in petto heeft.

Herman Cools. Ik heb de afgelopen jaren enorm genoten en veel geleerd. Voor mij was je een nieuw fenomeen dat ik mettertijd enorm ben gaan waarderen. Dank voor je vertrouwen en steun. Het artikel moet er nog komen, maar je ziet – als ik zeg dat iets er komt, dan komt het er.

Ineke, Peter, Mieke, Sofieke, Gerda, Derck, Erik, Pauline, Helma, Joost, Ellen, Karin, Jo, Else, Berthon, Pieter, en Marietta; Vrienden van het eerste uur, nog steeds uit Utrecht of van nog verder daarvoor, en inmiddels verspreid over Nederland. Contact hebben is de afgelopen jaren wel eens veel beperkter geweest dan we eigenlijk zouden willen. Vrienden niet spreken vanwege het afmaken van een proefschrift accepteren we, maar de duur ervan is mijns inziens aan een absoluut maximum gebonden. Daarom heb ik het klusje maar even afgemaakt. Dank voor jullie begrip en geduld; ik verheug me erop elkaar weer vaker te spreken!

Buren van ‘de witte huizen’. Jullie hebben het allemaal van nabij meegemaakt en zo je eigen conclusies getrokken. Ik geef je geen ongelijk! Dat jullie er zo bij betrokken waren de afgelopen jaren geeft al aan wat we opgebouwd hebben. Ik hoop dat de vriendschap nog heel lang mag duren.

Thea en Will, lieve schoonouders en schoonfamilie; Leon en Ursula, Marc en Dionne en de kinderen Robert, Marco en Twan. Dank voor jullie interesse, zorg en steun. Robert, wie weet komt je tante nu een keertje echt ‘on-line’.

Karin, van jou is het grafisch ontwerp van m’n proefschrift. Ik vind het helemaal te gek dat je dat voor me hebt willen doen. Ik heb een grenzeloze bewondering voor je doorzettingsvermogen en kijk uit naar al die goede tijden die er nog voor ons liggen.

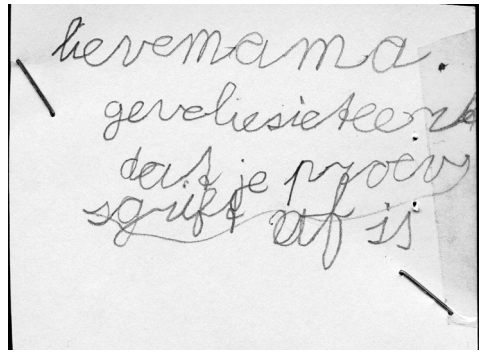
Renate, jij hebt ervoor gezorgd dat ik er vandaag ‘feestelijk en toch netjes’ uitzie. Ik voel me verwend en ben ontzettend blij dat je me ook voor deze gelegenheid hebt willen ‘aankleden’. En wie weet zitten er in de nabije toekomst meer gezamenlijke potten thee in dan we totnogtoe gered hebben.

Pappa en mamma, dank jullie wel voor het pakket aan genen dat jullie me hebben meegegeven waarmee ik al deze leuke dingen heb kunnen doen en voor de thuisbasis waar ik verder op heb kunnen bouwen. Mamma, je zou oorspronkelijk alleen oproepoma worden. Het zijn alleen wel erg veel woensdagen geworden. En bovendien zijn we er behoorlijk verslaafd aan geraakt. Ik wil jullie danken voor al jullie hulp, en bovenal jullie onvoorwaardelijke liefde.

Pascal, Nienke en Milet. Jullie weten er alles van, dat gepromoveer van die moeders. Jullie wil ik danken voor alles wat jullie aangehoord hebben, de gelegenheden die jullie geschapen hebben, de ondersteuning die jullie geboden hebben, of de afleiding waar jullie voor gezorgd hebben.
Madelon, het was een heerlijke tijd. Als ‘die promoverende moeders’ gingen we door dezelfde bergen en dalen. Het was leuk om elkaars werk zo goed te leren kennen, tegenslagen te relativeren en te ervaren dat er meer wegen zijn die naar Rome leiden. Ik wou dat ik een collage kon maken van alle omstandigheden waaronder we aan AsthmaCritic gewerkt hebben (vanuit de camper, ondersteboven in het

ziekenhuis, op ons buik tussen de bergen). Het is eigenlijk reuze jammer dat we nu allebei klaar zijn. Misschien kunnen we een nieuw project verzinnen?

Lieve Eric, Titia, Otto, en Sascha. Het grote onderzoek is klaar. Nu echt jongens! Het feest kan beginnen, geen zondagen meer boven op zolder, niet alle avonden meer bezet. Eric, als het te veel wennen is zeg je het maar, maar promoveren doe ik maar één keer!



CURRICULUM VITAE

Manon Marguerite Kuilboer was born on June 6th, 1964 in Amsterdam, The Netherlands. She received her undergraduate education at the Corderius College in Amersfoort (gymnasium-B) from 1976-1982.

In 1982 she started studying Medicine in Utrecht, which she completed in 1990. Meanwhile, she worked for the Department of Development and Research for Medical Education (O.O.&O.) at the University of Utrecht (1984-1987) and started taking courses in computer science at the Department of Information Science of Mathematics in Utrecht (1983, 1986-1987). From 1990 to 1991 she worked as a medical researcher for the 'Wilhelmina Kinderziekenhuis' (Wilhelmina Children's Hospital).

Between 1991 and 1993 she completed the study Master of Science in Medical Information Sciences at Stanford University School of Medicine, in the United States of America.

In December 1993 she started the research reported in this thesis. From 1999-2002, she worked for the Department of General Practice and Nursing Home Care at the Leiden University Medical Center, to develop a Nursing Home Information system.

In September 2002, she started working with 'Microbais Automatisering BV' in Amsterdam as Product Manager Medical Systems.

She lives in Breda with her husband Eric Broers and their three children: Titia, Otto, and Sascha.


Ac
Ba
Pa
Da
De
AC
fo