IMPROVING EVIDENCE-BASED GENERAL PRACTICE

Marlous Frédérique Kortekaas

Improving evidence-based general practice

Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, PhD thesis, with a summary in Dutch. University Utrecht, Faculty of Medicine.

ISBN:	978-94-6182-701-2
Author:	Marlous F. Kortekaas
Cover:	Painting: Gabriëlle Kortekaas Lay-out: Pieter Koningsveld
Lay-out & print:	Off Page, Amsterdam

IMPROVING EVIDENCE-BASED GENERAL PRACTICE

Verbetering van wetenschappelijk onderbouwd handelen door de huisarts
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad
van doctor aan de Universiteit Utrecht op gezag
van de rector magnificus, prof. dr. G.J. van der Zwaan,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op donderdag
29 september 2016 des ochtends te 10.30 uur

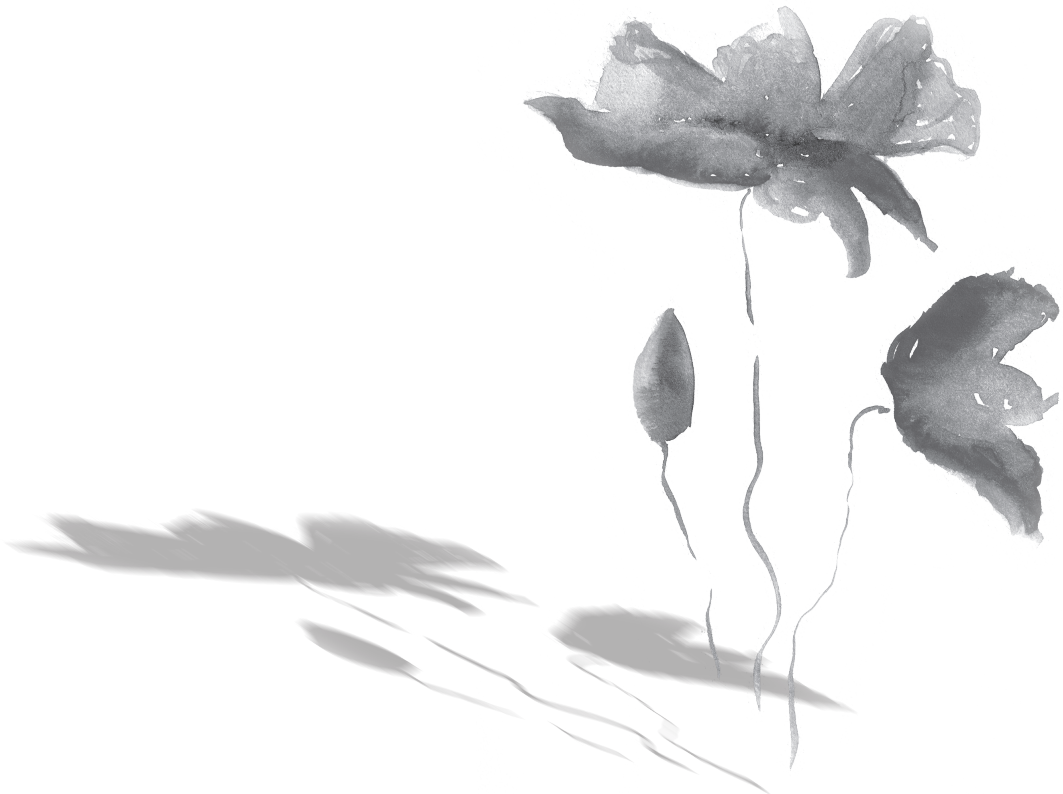
door

Marlous Frédérique Kortekaas
geboren op 13 maart 1983
te Wassenaar

Promotoren:	Prof. dr. N.J. de Wit Prof. dr. A.W. Hoes
Copromotor:	Dr. M.E.L. Bartelink

The printing of this thesis was financially supported by SBOH, employer of GP trainees in the Netherlands

TABLE OF CONTENTS



CHAPTER 1

General introduction

"Wat je niet begrijpt, moet je bewonderen."
Piet den Ouden

A clinical query about a patient with Bell's palsy... 25 years ago

At the beginning of the 1990s, Mr. Jansen, a male patient of 27 years old, enters the waiting room for an appointment with his general practitioner (GP). Since that morning he suddenly can't close his mouth and lost control of the muscles on the left side of his face. After five minutes of waiting, the GP trainee takes him to the consultation room. The trainee asks about the patient's complaints, and performs a physical examination. After a while the GP trainee tells the patient he believes it is Bell's palsy, explains what it is, and says that he is going to discuss the best treatment option with his GP supervisor.

The trainee walks to his supervisor, who has been working as a GP for more than twenty years. He tells his supervisor that he told the patient he wasn't sure about the best treatment option and that he wanted his supervisor's opinion. The GP supervisor urges the trainee never again to tell a patient that he is insecure on the best way to manage a patient. He further tells his trainee that an ENT specialist, a friend since medical school, told him that in case of Bell's palsy it is best to prescribe both antiviral medication and corticosteroids; he always treats patients this way, and he has never seen any persisting paresis.

The GP trainee goes back to the patient and prescribes both the antiviral medication and corticosteroids. He explains to the patient that he is absolutely sure this is the best treatment option, as both his supervisor and an ENT specialist agree.

He asks the patient to come back in one week, but to contact him earlier in case the paresis worsens.

A clinical query about a patient with Bell's palsy... 2016

Mrs. Jansen, 51 years, consults a third year trainee, because she suddenly can't close her mouth anymore and lost control of the muscles on the right side of her face. The trainee asks about her complaints, performs a physical examination, and concludes that the patient suffers from Bell's palsy. The patient says her husband encountered the same symptoms 25 years ago and received both antiviral medication and corticosteroids from their former GP. She asks if the GP trainee could prescribe the same medication, since it had helped her husband so well back then.

The GP trainee listens to the patient's request, and explains that she wants to search for the latest evidence on this topic. The GP trainee tells the patient there is a clinical practice guideline (CPG) from the Dutch College of GPs on "peripheral facial paralysis", but that Bell's palsy is so rare that she does not know the recommended therapy by heart and she needs to check.

With the patient in front of her she goes to the website of the Dutch College of GPs and opens the digital version of this CPG to retrieve information on the treatment options for Bell's palsy. The CPG recommends prescribing corticosteroids 25 mgs twice a day as soon as possible after the diagnosis and refers to a randomized controlled trial from 2007 as the source of the evidence underlying this recommendation.(1) It also states that there is no indication to prescribe antiviral medication, since this will not improve the recovery process. The latter information is based on two meta-analyses from 2009.(1,2)

The GP trainee explains to the patient that she will prescribe corticosteroids again, but that she will not prescribe antiviral medication, since this will not improve recovery. The patient is somewhat surprised not to receive antivirals, but agrees after the GP trainee's explanation.

The GP trainee refers the patient to the patient website of the Dutch College of General Practitioners, "www.thuisarts.nl", where she can find more information on her disease, and advises to make a return appointment in two days.

IMPROVING EVIDENCE-BASED GENERAL PRACTICE

Development of general practice as a medical specialty

General practitioners (GPs) play an important role as gatekeeper in the health care system in the Netherlands. Patients are usually first seen in general practice, and will be seen by specialists only after referral. Every Dutch resident has to register in one general practice. Health insurance is mandatory and covers a standard benefit package including general practice care.(3) The 9000 GPs in the Netherlands address more than 90% of the health problems on a budget of less than 5% of the national healthcare expenditure.(4) General practices are easily accessible during office hours and collaborate in out-of-hour services.(3) According to annual surveys, the overall satisfaction with general practice care is consistently high, with a mean score of 7.9 (on a scale of 1-10).(5)

General practice has evolved considerably over the past decades. In 1956 the Dutch College of GPs was founded, aiming to improve the quality of care provided by GPs.(6) Continuous, integral, and personal care was labelled as the cornerstone of general practice.(7) It was only since 1974, however, that general practice was acknowledged as a medical specialty and that a mandatory GP registry with repeated accreditation was installed. In 1974 the GP specialty training was introduced; initially with a one-year program, which was prolonged to two years in 1988, and three years in 1994. Gradually general practice was embedded in the universities in the Netherlands, but it was not until 1966 that Jan van Es was appointed at Utrecht University as the first professor of general practice in the Netherlands.(6)

Research and clinical practice guidelines

Scientific research is of great importance for the professional development of medical disciplines.(8) Results derived from research undertaken in the hospital setting cannot uncritically be applied in general practice, because of major differences in patient characteristics. The general practice population can be characterized by a four-category model; the "the truly healthy", "the imagined ill", "the truly ill", and "the imagined healthy" (9), whereas the hospital population typically consists of the second and third category merely. This results in different a-priori chances of diseases in both groups, while the patients' prognosis is also likely to differ between the settings. Consequently, extrapolation of the findings from one setting to the other is not self-evident, and often general practice based studies are required to guide patient management in general practice.(10)

Over the past decades, the volume of research in general practice has been expanding worldwide. A bibliometric analysis from 2013 demonstrated that the Netherlands are among the top ranking countries in clinical research in general practice, both in quality and in quantity.(10) As the budgets for scientific research are being reduced, research topics need to be prioritised.(11) Efficient use of resources for research in general practice requires that future projects should address clinically

relevant questions, and focus on problems that are frequently encountered in everyday clinical practice. This so-called practice based research in general practice is essential to fill the knowledge gaps in clinical practice guidelines (CPG), and provides the fundamentals for evidence-based medicine (EBM).(12)

CPGs play an important role in the quality of care program of the Dutch College of GPs in the Netherlands. In 1989 the first national CPG for GPs was published, on general practice management of Diabetes.(13) Presently there are more than 100 highly appreciated evidence-based CPGs developed by the College, resulting from a structured process based on assessment of research evidence, and expert opinion.(14) These CPGs provide GPs with evidence-based recommendations for daily clinical practice on a wide variety of common clinical topics. As CPGs do not cover all clinical topics encountered in daily clinical practice, GPs have to retrieve the latest evidence themselves from other resources. This requires training in skills to enable them to find, interpret and use relevant evidence.

Evidence-based medicine

The term evidence–based medicine (EBM) originates from the work of Pierre Louis in Paris in the 19th century, but was introduced as the cornerstone of today's medical practice only 20 years ago. EBM is defined as the integration of clinical expertise, patient values and the best available clinical evidence in daily clinical practice.(12) The process of practising EBM is summarised in five steps, as stated in the Sicily Statement; ask, acquire, appraise, apply and assess:

Ask:	Translation of a clinical problem into an answerable question.
Acquire:	Efficient search for the best evidence to fit the question.
Appraise:	Appraising the evidence retrieved regarding methodological quality and the applicability to individual patients.
Apply:	Making a decision in respect of the best available evidence, clinical expertise, and patient values.
Assess:	Evaluation of the quality of this process.

An essential step in practising EBM is 'acquire'; which requires the skill to search for evidence-based answers to clinical queries encountered in daily clinical practice. Physicians frequently have questions about the most appropriate patient management in daily practice that they cannot answer immediately. A review by Del Fiol et al (2014) showed that doctors have on average two questions per every three patient consultations, that they search for answers for half of these questions, and that this search results in an answer in 80% of these.(15) These queries do have an impact on clinical management: evidence shows that in half of the consultations in which doctors search for evidence this changes patient management.(12,16) Trainees probably differ from experienced doctors in their information needs and seeking behaviour, as they

have limited clinical experience, are better trained in practising EBM, and are more accustomed to using internet based resources. There are different ways to deal with queries and different ways to search for answers. The evidence required to answer queries can be ranked based on the quality of the available evidence in the so-called EBM pyramid.(17) According to the hierarchy of information resources for evidence, doctors can use 1) guidelines/pre-appraised resources, 2) unfiltered information sources, and/or 3) expert opinion or background information. To apply EBM in daily clinical practice training is needed to gain the essential knowledge and skills to search for evidence-based answers to clinical queries.

Training in evidence-based medicine

To qualify as general practitioner, medical graduates in the Netherlands follow a three-year GP specialty training program. This training combines training in clinical practice (four days a week) with a one-day educational program in a group of 12 at the training institute. The training is competency based, in accordance with the CanMEDS Physician Competency Framework. This describes the knowledge, skills and abilities that specialist physicians need.(18) One of the core CanMEDS competencies is to work according to the principles of EBM.(18) The main goal of EBM training is to teach GP trainees to integrate EBM into their clinical practice and apply it in daily patient care. With the rapidly expanding medical knowledge it is important that (future) GPs can keep up-to-date with current evidence. By learning to combine available evidence with clinical expertise and patient values, trainees are equipped to deliver adequate patient care.

Despite its undisputed importance, evidence for the optimal way to teach practitioners to apply EBM in daily clinical practice is limited.(19,20) EBM is often trained in stand-alone courses, taking place in classrooms, and with content not necessarily related to clinical practice.(21) Typically, emphasis is on research skills, clinical epidemiology, and statistics, whereas the use of evidence in daily clinical practice, while the integration of evidence with clinical expertise and patient values, receives less attention.(22) An alternative approach to train EBM is through so-called integrated EBM training; this implies teaching EBM in the clinical context, and starting teaching sessions from clinical dilemmas encountered by the trainee in daily practice.(21) Both stand-alone and integrated EBM training improve EBM knowledge. According to a review from 2004, the skills, attitude, and behaviour required for evidence-based practice seem to improve only when integrated EBM training is applied.(21) Authors of that review conclude that "the teaching of EBM should be moved from classrooms to clinical practice to achieve substantial improvements in outcomes".(21) However, the effectiveness of stand-alone and integrated EBM training was never directly compared in a (randomized) controlled trial, nor were the effects on EBM behaviour of former trainees assessed in daily clinical practice after graduation.(21,23) Moreover, most research on the effects of EBM training was performed in trainees in a hospital-based specialty, and it is unknown whether integrated EBM training also improves EBM behaviour in GP trainees.

Effects of training in evidence-based medicine

EBM training should improve knowledge, attitude, skills, and – most importantly – EBM behaviour.(21,24,25) Many instruments have been developed to measure the effects of EBM training on one or more of these outcomes.(24–26) To be able to apply EBM in clinical practice it is essential to have basic knowledge on the principles of clinical epidemiology. Frequently used methods to measure this knowledge are questionnaires. Most of the current questionnaires (the Fresno, the Berlin, and the ACE tool) do not focus on knowledge relevant for clinical practice, are time-consuming to score, or focus on therapeutic issues only, whereas many reasons for encounter in daily practice regard diagnostic or prognostic topics.(27–29)

In addition to EBM knowledge, the effects of EBM training should also be assessed on EBM behaviour.(26) To date, however, there is no agreement on the optimal way to measure EBM behaviour.(24) It is suggested that it is best assessed by monitoring clinical management decisions (30), but there are no validated instruments for this. (24,25) Since an essential skill of EBM practice is the adequate search for evidence in daily clinical practice (26), assessment of the information needs and seeking behaviour of trainees could be used as a proxy for EBM behaviour. Information on the type and frequency of clinical queries and on the approach they take to pursue answers would be instrumental to optimize EBM training. Another way to measure the effects of EBM training on EBM behaviour in daily practice is to assess adherence to CPGs. Rational guideline adherence means clinical management in accordance with key recommendations in professional guidelines, while allowing motivated deviation in individual patients. As the clinical content of the GP specialty training in the Netherlands is based on the CPGs developed by the College, monitoring CPG adherence could be a way to assess EBM behaviour during professional training.(31)

OBJECTIVE AND OUTLINE OF THIS THESIS

The aim of this thesis is to gain more knowledge on how to improve the practice of EBM in general practice. We first focused on the match between the current evidence base in general practice and the knowledge needs from clinical practice (practice-based evidence). Next, we focused on training in evidence-based practice, by studying the information needs and seeking behaviour of GP trainees, and comparing the effects of two different EBM training programs, integrated vs. stand-alone, on outcomes relevant to daily clinical practice measured with reliable and valid instruments.

In ***chapter 2*** we describe a bibliometric analysis of all internationally published RCTs within the field of general practice in the period covering 1990 to 2010, and assessed if the distribution of research topics matched with clinical practice in the Netherlands. In ***chapter 3*** we describe all on-going clinical research projects in the field of general practice in the Netherlands. We constructed a nationwide database of on-going research projects in the field of general practice in the Netherlands, and again assessed if the distribution of research topics matched with clinical practice. The aim was to enable a better match between future research and the current needs from general practice. In the next chapters we focussed on EBM training. In ***chapter 4*** a systematic review on information needs and seeking behaviour of trainees is presented. In ***chapter 5***, we assessed how and how often GP trainees search for answers to clinical queries they encounter in daily clinical practice. In chapters 6 and 7 we report two instruments to measure the effect of EBM training on EBM knowledge and EBM behaviour. In ***chapter 6,*** the development and validation of the Utrecht questionnaire on knowledge on Clinical epidemiology for Evidence-based Practice (U-CEP), an instrument to assess the effect of EBM training is presented. In ***chapter 7*** we report on the development and validation of an instrument that assesses guideline adherence of GP trainees in clinical practice, while allowing motivated deviation in individual patients. Working in accordance with guidelines is considered as one of the key indicators of practising EBM, and monitoring guideline adherence is a way to assess EBM behaviour during professional training.(31) In ***chapter 8*** we report the results of a cluster randomized controlled trial (RCT) among GP trainees assessing the effects of an integrated EBM training program, compared to stand-alone EBM training. We assessed the effects by measuring knowledge on clinical epidemiology, EBM attitude, and EBM behaviour, both during training as well as during daily clinical practice after graduation. Finally, in ***chapter 9*** we discuss the main findings, relate them to existing knowledge and identify key challenges for successful EBM training in the general practice setting.

REFERENCES

1 Sullivan FM, Swan IRC, Donnan PT, Morrison JM, Smith BH, McKinstry B, et al. Early treatment with prednisolone or acyclovir in Bell's palsy. N Engl J Med [Internet]. 2007 Oct 18;357(16):1598–607.

2 de Almeida JR, Al Khabori M, Guyatt GH, Witterick IJ, Lin VYW, Nedzelski JM, et al. Combined corticosteroid and antiviral treatment for Bell palsy: a systematic review and meta-analysis. JAMA. 2009 Sep 2;302(9):985–93.

3 Faber MJ, Burgers JS, Westert GP. A sustainable primary care system: lessons from the Netherlands. J Ambul Care Manage [Internet]. Jan;35(3):174–81.

4 LHV [Internet]. [cited 2016 May 26].

5 Peters S. Nederlanders blij met hun huisarts. LHV - de dokter. 2015 Dec;8–11.

6 Van Es JC. Een halve eeuw huisartsgeneeskunde. Van ambacht naar professie. [Internet]. Houten; 2006.

7 Huygen FJA. NHG - Quo Vadis. Huisarts Wet. 1959;2:187–91.

8 Hummers-Pradier E, Beyer M, Chevallier P, Eilat-Tsanani S, Lionis C, Peremans L, et al. The Research Agenda for General Practice/ Family Medicine and Primary Health Care in Europe. Part 1. Background and methodology. Eur J Gen Pract [Internet]. 2009 Dec;15(4):243–50.

9 Hetlevik I. Evidence-based medicine in general practice: a hindrance to optimal medical care? Scand J Prim Health Care [Internet]. 2004 Sep;22(3):136–40.

10 Glanville J, Kendrick T, McNally R, Campbell J, Hobbs FDR. Research output on primary care in Australia, Canada, Germany, the Netherlands, the United Kingdom, and the United States: bibliometric analysis. BMJ. 2011 Jan;342:d1028.

11 Dorsey ER, de Roulet J, Thompson JP, Reminick JI, Thai A, White-Stellato Z, et al. Funding of US biomedical research, 2003-2008. JAMA [Internet]. 2010 Jan 13;303(2):137–43.

12 Sackett DL, Rosenberg WMC, Gray J a M, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. 1996. BMJ. 1996 Feb;312(71).

13 Cromme P, Mulder JD, Rutten GEHM, Zuidweg J TS. Nederlands Huisartsen Genootschap. Diabetes mellitus type II. Huisarts Wet. 1989;32:15–8.

14 Handleiding. Ontwikkeling van NHG-standaarden. 2.0 ed. Utrecht: Nederlands Huisartsen Genootschap, afdeling richtlijnontwikkeling en wetenschap; 2015. 7 p.

15 Del Fiol G, Workman TE, Gorman PN. Clinical questions raised by clinicians at the point of care: a systematic review. JAMA Intern Med. 2014;174(5):710–8.

16 Crowley SD, Owens T a, Schardt CM, Wardell SI, Peterson J, Garrison S, et al. A Web-based compendium of clinical questions and medical evidence to educate internal medicine residents. Acad Med. 2003;78(3):270–4.

17 The Sladen Library and Center for Health Information Resources. The EBM Pyramid of Evidence. [Internet]. [cited 2016 May 26].

18 Frank JR. The CanMEDS 2005 physician competency framework. Better standards. Better physicians. Better care. Ottawa R Coll Physicians Surg Canada. 2005;

19 Ilic D, Maloney S. Methods of teaching medical trainees evidence-based medicine: A systematic review. Med Educ. 2014;48(2):124–35.

20 Maggio LA, Tannery NH, Chen HC. Evidence-Based Medicine Training in Undergraduate Medical Education: A Review and Critique of the Literature Published 2006–2011. 2013;88(7):1–7.

21 Coomarasamy A, Khan KS. What is the evidence that postgraduate teaching in evidence based medicine changes anything? A systematic review. BMJ. 2004;329(7473):1017.

22 Kelly MP, Heath I, Howick J, Greenhalgh T. The importance of values in evidence-based medicine. BMC Med Ethics. 2015 Oct 12;16(1):69.

23 Taylor R, Reeves B, Ewings P, Binns S, Keast J, Mears R. A systematic review of the effectiveness of critical appraisal skills training for clinicians. Med Educ. 2000;34(2):120–5.

24 Oude Rengerink K, Zwolsman SE, Ubbink DT, Mol BWJ, van Dijk N, Vermeulen H. Tools to assess evidence-based practice behaviour among healthcare professionals. Evid Based Med. 2013 Aug;18(4):129–38.

25 Shaneyfelt T, Baum KD, Bell D, Feldstein D, Houston TK, Kaatz S, et al. Instruments for evaluating education in evidence-based practice: a systematic review. JAMA. 2006 Sep 6;296(9):1116–27.

26 Dawes M, Summerskill W, Glasziou P, Cartabellotta A, Martin J, Hopayian K, et al. Sicily statement on evidence-based practice. BMC Med Educ. 2005 Jan 5;5(1):1.

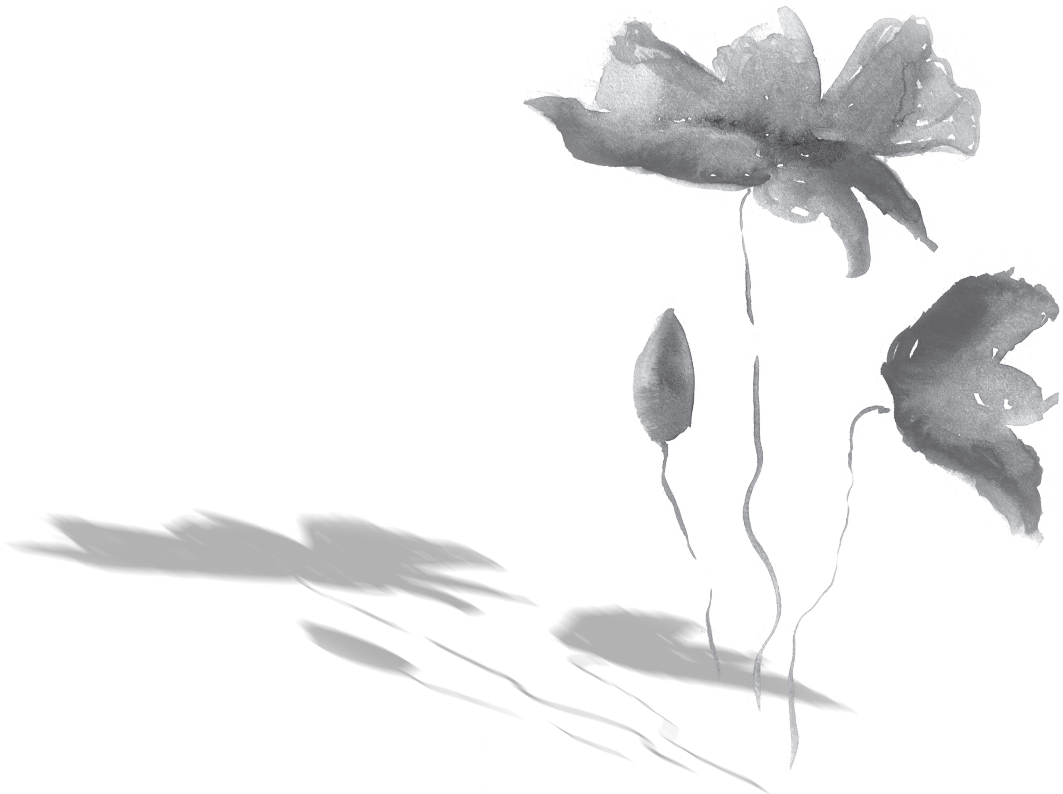
27 Fritsche L, Greenhalgh T, Falck-Ytter Y, Neumayer H-H, Kunz R. Do short courses in evidence based medicine improve knowledge and skills? Validation of Berlin questionnaire and before and after study of courses in evidence based medicine. BMJ. England; 2002 Dec;325(7376):1338–41.

28 Ramos KD, Schafer S, Tracz SM. Validation of the Fresno test of competence in evidence based medicine. BMJ. England; 2003 Feb;326(7384):319–21.

29 Ilic D, Nordin R Bin, Glasziou P, Tilson JK, Villanueva E. Development and validation of the ACE tool: assessing medical trainees' competency in evidence based medicine. BMC Med Educ [Internet]. 2014 Jan;14:114.

30 Tilson JK, Kaplan SL, Harris JL, Hutchinson A, Ilic D, Niederman R, et al. Sicily statement on classification and development of evidence-based practice learning assessment tools. BMC Med Educ [Internet]. BioMed Central Ltd; 2011;11(1):78.

31 de Belvis a. G, Pelone F, Biasco a., Ricciardi W, Volpe M. Can primary care professionals' adherence to Evidence Based Medicine tools improve quality of care in Type 2 diabetes mellitus? A systematic review. Diabetes Res Clin Pract. 2009;85(2):119–31.



CHAPTER 2

Randomized controlled trials in general practice worldwide: analysis of themes, quality and match with Dutch clinical practice in the last 20 years

Based on: Kortekaas MF, Meijer A, Van de Pol AC, De Wit NJ. RCT's in de huisartsgeneeskunde, 1990-2010. Huisarts Wet 2012;55(11):486-91

Kortekaas MF, Meijer A, Van de Pol AC, De Wit NJ

"Als je iets weet, maar het dringt niet tot je door, dan weet je eigenlijk niets."
Toon Hermans

ABSTRACT

Introduction

Research output in the area of primary care medicine has developed substantially over time, but publication volume varies over the different research themes. The goal of our bibliometric analysis was to assess the quantity, themes and quality of randomized controlled trials (RCTs) in general practice that have been conducted in the past and to assess matching of research themes with daily clinical practice.

Methods

Published general practice RCTs were identified through a systematic search in the Medline database (1990-2010). We included articles using predefined inclusion criteria. Publication date, country of origin, study design, journal, and impact factor were recorded. We categorized research themes of studies using International Classification of Primary Care (ICPC) codes.

Results

A total of 1935 RCTs were included. The amount of RCTs published each year has increased roughly from 50 in 1991 to 170 in 2009. The United States (28%), the United Kingdom (26%) and the Netherlands (10%) conducted the majority of studies. The median impact factor of publications was variable, ranging from approximately 2.2 to 3.7 over the years. Most RCTs addressed the field of psychiatry (code P; 28%), circulatory (code K, 13%), general (code A, 13%), endocrinologic / metabolic (code T, 11%) and respiratory complaints (code R, 10%). The distribution of research themes did not adequately match morbidity as presented in general practice.

Discussion

Although the number of RCTs in general practice continued to rise over the past 20 years, the match between research fields and clinical practice could be improved.

INTRODUCTION

Scientific research is of great importance for the professional development of medical disciplines.(1,2) Research in the field of general practice (GP) has undergone rapid development over the past thirty years, in quantity as well as quality.(3-5) A recent analysis shows that Dutch general practice is doing well compared to other countries with regard to the total research output and is even leading in return on investment for research funds, as based on a random sample (15%) of 82,169 general practice publications from the United Kingdom, the United States, Germany, the Netherlands, Canada and Australia between 2001 and 2007.(6) This analysis showed the number, authors and impact of the research output, but did not take into account the method and subject of the research. Earlier research showed that the number of publications in the Netherlands is increasing and covering a broad spectrum of disorders, but the topics studied do not correspond well with the diseases most prevalent in general practice.(7)

Apart from the research themes and the match with clinical practice, the quality of research is also important. In general, the randomized controlled trial (RCT) is regarded as the optimum design for therapeutic, clinical research.(8) In order to find out how RCT volume and themes within general practice have developed in recent history, we conducted a bibliometric analysis of internationally published RCTs within the field of general practice over the period 1990-2010. We analyzed the number and quantity of these publications, and assessed if the distribution of research topics matched with clinical practice in the Netherlands.

METHODS

Search strategy

In the MEDLINE database we searched for articles published within the period between January 1990 until March 2010 (search date 23/05/2010) using the terms 'general practice' and 'primary health care'. The exact text of the search was as follows: ("general practice"[MeSH Terms] OR "primary health care"[MeSH Terms] OR "general practice"[Title/Abstract] OR "primary health care"[Title/Abstract] OR "family medicine"[Title/Abstract]) AND English[lang] AND "1990/01/01"[PDAT]: "2010/03/24"[PDAT].(9) With the help of PubMed filters 'Clinical Trial' and 'Randomized Controlled Trial' we selected RCTs out of the search results.(7)

Selection of RCTs

We included articles using predefined inclusion criteria: (1) original RCT; (2) clinical subject related to general practice; (3) research in the domain of general practice; (4) therapeutic research; (5) abstract written in English and/or Dutch, and (6) abstract and/or full-text available in the Netherlands. The terminology used in an article was leading. Title and summary usually contained sufficient information; in only ten cases it

was necessary to assess full-text. The selection was carried out by AM, assisted where necessary by MK (in approximately 2.5% of the articles). Disagreement was resolved by seeking consensus or, in case of persistence, a third investigator (NdW) made a final decision. To check the quality of this selection, we compared our selection to that of various meta-analyses. With the use of the PubMed limit 'Meta-Analysis' we found 109 meta-analyses, from which we took a random sample of five. The RCTs in these five meta-analyses turned out to be present in our own selection.

Data abstraction

Publication date, country of origin, study design, journal, and impact factor were recorded. We used the impact factor of the journal for 2010 (found through Web of Knowledge) as an indicator of the quality of the publication.(10) In order to compare the number of publications to country of origin, we also gathered information from all the countries about their gross domestic product and population size.(6,11) Finally, we categorized research themes of studies using International Classification of Primary Care (ICPC) codes.(12) The articles that could not be categorized this way, were categorized as 'other'.

Data analysis

We calculated the research output per country per year, with a correction based on gross domestic product and population size for countries with two or more publications in total. For each year we assessed the median impact factor of the journals in which the RCTs were published. Then we assessed to which degree the research themes matched daily clinical practice by comparing the themes of the publications with the distribution of clinical themes in the Dutch general practice.(12)

RESULTS

Selected studies

Origin and number of publications

A total of 82,577 potentially relevant studies were retrieved of which 6064 were RCTs. After applying the inclusion criteria, a final 1935 of relevant RCTs remained to be assessed (figure 1). As figure 2 shows, the number of published RCTs regarding the field of general practice continually grew over the researched period, internationally as well as nationally. In 1990 we found 48 RCTs worldwide. That number increased to 161 in 2009 – which is 3.4 times as many. For the Netherlands we even found that the number had increased fivefold, from 5 RCTs in 1990 to 25 in 2009. Most RCTs were published by institutions in the United States (28.4%), the United Kingdom (26.4%) and the Netherlands (10.0%). Together the institutions from these countries produced almost two thirds of the total (64.8%). In relation to the gross domestic product, the Netherlands had the biggest yearly output of RCTs, followed by New Zealand and the

Figure 1. Flowchart search strategy.

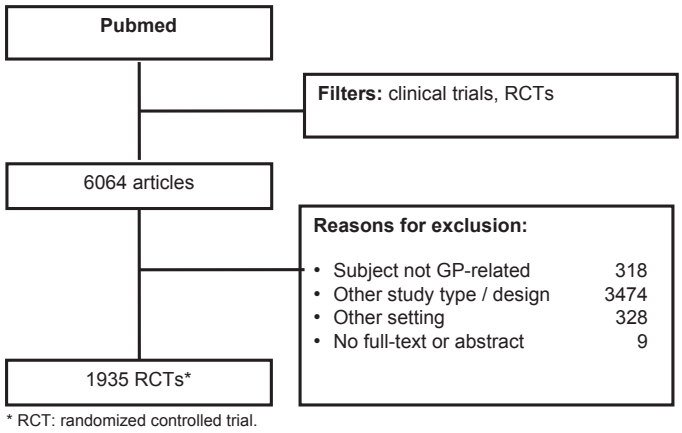
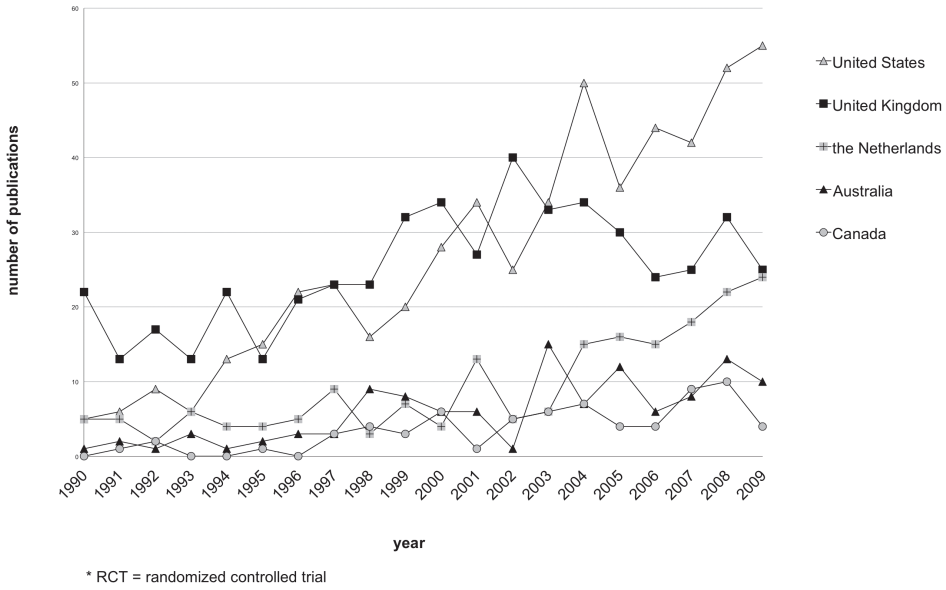


Figure 2. Top 5 countries: number of published RCTs* within the field of general practice per year .

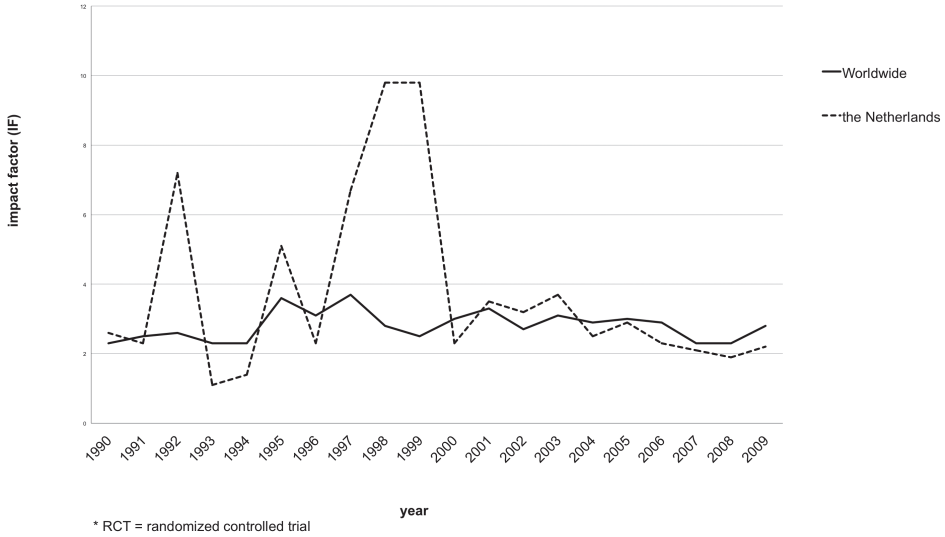


United Kingdom. Measured against population size, the yearly output was the largest in the Netherlands, Norway and Denmark.

Quality of publications

The median impact factor of the journals in which internationally published RCTs were published between 1990 and 2010 was between 2.2 and 3.7, for RCTs from the

Figure 3. Median impact factor of RCTs* per year in the field of general practice.



* RCT = randomized controlled trial

Netherlands this varied between 1.1 and 9.8 (figure 3). At the time of the analyses the impact factor of 200 RCTs (10%) was unknown. The RCTs that we found were published in a broad number of journals. The three journals publishing most RCTs were the *British Medical Journal* (7.6%), the *British Journal of General Practice* (5.1%) and *Family Practice* (2.6%). However, the largest part by far (79.1%) was published in journals that were not in the top ten. The top three of journals contained research

Table 1. Number of publications per country and per journal.

	IF$	United States	United Kingdom	Scandinavia*	Netherlands	Other	Total per journal
BMJ**	12,8	6	100	7	18	17	**148**
Br J Gen Pract	2,3	1	70	5	19	4	**99**
Fam Pract	1,63	0	18	12	6	14	**50**
JAMA**	28,9	27	1	0	0	2	**30**
J Fam Pract	1,4	16	1	1	2	1	**21**
Lancet**	28,4	0	13	0	4	2	**19**
Ann Intern Med	17,5	13	1	0	1	2	**17**
Ann Fam Med	3,5	13	0	0	1	1	**15**
NEJM**	50,0	1	0	0	0	0	**1**
Fam Med	1,6	3	0	0	0	2	**5**
Other		469	307	176	143	435	**1530**
Total per country		**549**	**511**	**201**	**194**	**480**	**1935**

IF = impact factor. * Scandinavia: Sweden, Norway, Finland, and Denmark. **Highest impact factors

from the United Kingdom and the Netherlands primarily; of the 30 publications in the journal that ranked fourth (JAMA), 27 (90%) came from the United States (table 1).

Subjects of the RCTs

Table 2 shows the distribution of thematic RCTs on the ICPC chapters. In total, 54 international RCTs (2.8%) could not be allocated and ended up in the section 'other'. Most of the research over the total control period concerned ICPC chapter P, mental health problems (n = 547, 28.3%), followed by chapter K, circulatory tract (n = 249, 12.9%), chapter A, general and non-specified (n = 243, 12.6%), chapter T, endocrine glands / metabolism / nutrition (n = 213, 11.0%), and chapter R, respiratory tract (n = 188, 9.7%).

During the period 1990 up until 2009 the biggest absolute increase in the number of publications (from 12 in 1990 to 47 in 2009; an increase of nearly 300%) is apparent in ICPC chapter P, psychiatric disorders. Other research subjects that have increased relatively steeply are general disorders (from 1 to 20, factor 20), endocrine / metabolic disorders (from 2 to 25, more than factor 12) and the circulatory tract (from 7 to 20, almost threefold). Apart from that, it is noticeable that fewer RCTs performed between 1990 and 2009 were related to the urinary tract (from 10 in 1990 to four in 2009, more than halved). Upon further categorization within ICPC chapters, 189 studies could

Table 2. Topics of general practice related RCTs from the Netherlands between 1990-2010.

ICPC chapter		All RCTs		RCTs from the Netherlands		Prevalence Netherlands
A	General	243	(12.6%)	19	(9.8%)	4.9%
B	Blood	13	(0.7%)	0	(0.0%)	0.9%
D	Digestive tract	76	(3.9%)	8	(4.1%)	7.3%
F	Eyes	5	(0.3%)	0	(0.0%)	3.2%
H	Ears	23	(1.2%)	3	(1.5%)	4.6%
K	Circulatory tract	249	(12.9%)	25	(12.9%)	7.0%
L	Musculoskeletal system	152	(7.9%)	31	(16.0%)	16.2%
N	Nervous system	32	(1.7%)	1	(0.5%)	3.0%
P	Mental health problems	547	(28.3%)	34	(17.5%)	5.1%
R	Respiratory tract	188	(9.7%)	33	(17.0%)	13.0%
S	Skin and subcutaneous tissue	17	(0.9%)	1	(0.5%)	14.7%
T	Endocrine glands	213	(11.0%)	21	(10.8%)	3.0%
U	Urinary tract	41	(2.1%)	10	(5.2%)	2.7%
W	Pregnancy	18	(0.9%)	0	(0.0%)	4.0%
X	Female genitals	43	(2.2%)	1	(0.5%)	7.3%
Y	Male genitals	11	(0.6%)	0	(0.0%)	2.2%
Z	Social problems	10	(0.5%)	1	(0.5%)	1.1%
	Other	54	(2.8%)	6	(3.1%)	0.0%
	Total	1935	(100%)	194	(100%)	(100%)

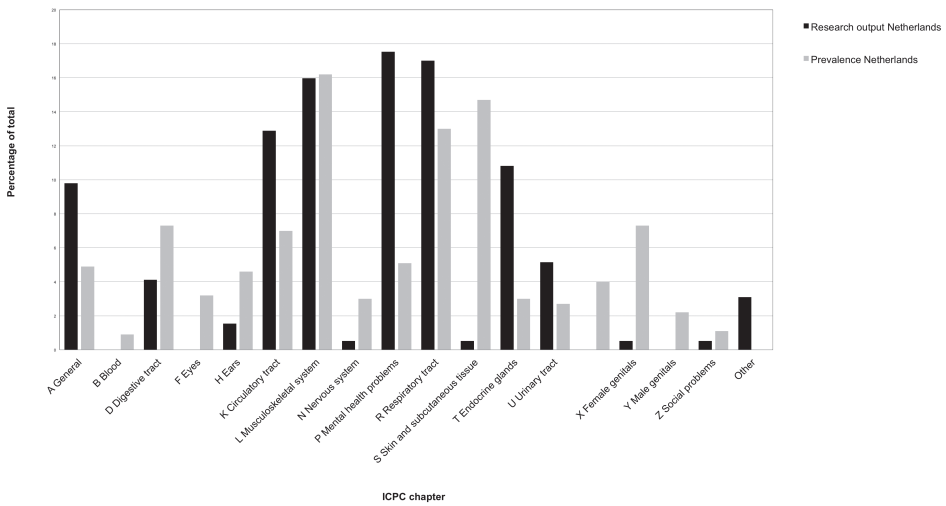
not be classified, and a total of 243 studies (12.6%) ended up in the category 'other'. Chronic diseases have been studied most, namely depression (code P76, 12.2%), hypertension (code K86, 7.1%), diabetes mellitus (code T90, 6.6%), asthma (code R96, 4.3 %) and tobacco abuse (code P17, 3.9%).

RCTs performed in the field of general practice in the Netherlands addressed psychological problems (n = 34, 17.5% of total), respiratory tract (n = 33, 17.0%), musculoskeletal system (n = 31, 16.0%) and circulatory tract (n = 25, 12.9%) primarily. Frequently, chronic diseases, such as diabetes mellitus (code T90, n = 18, 9.3% of the total), depression (code P76, n = 15, 7.7%), and asthma (code R96, n = 12, 6.2%) were addressed.

How research themes match daily practice

According to the Second National Study GP consultations in the Netherlands primarily concern the musculoskeletal system (16.2%), the skin and subcutaneous tissue (14.7%), and the respiratory tract (13.0%).(12) The research topics, however, have a completely different ranking (figure 4). In particular the discrepancy in the area of mental health problems stands out: 17.5% of RCTs versus 5.1% of the consultations. The amount of research that is done on common diseases, circulatory tract, respiratory tract and endocrine / metabolic disorders is also comparatively high. The percentage of dermatological issues researched is much lower than the frequency they are encountered in GP: 0.5% of RCTs versus 14.7% of the consultations. Moreover, clinical questions in the field of reproductive organs, pregnancy, eyes and the digestive tract are researched relatively less frequently. In the Netherlands solely research on the musculoskeletal

Figure 4. Comparison of research topics of RCTs within the field of general practice from the Netherlands (between 1990 and 2010) with prevalence of reasons for encounter (ICPC chapters) in the Netherlands*.



system corresponds with the frequency of the complaints encountered in GP, with 16.0% and 16.2% respectively. The five most common disorders in GP are K86 (essential hypertension), R74 (acute upper respiratory tract infection), S74 (dermatomycoses), R05 (cough), and S88 (eczema).(12) The five disorders most frequently being the topic addressed in an RCT are T90 (diabetes mellitus), P76 (depression), R96 (asthma), K86 (hypertension), K99 (other diseases circulatory tract) and L03 (lower back pain).

DISCUSSION

Over the past twenty years the number of GP related RCTs published worldwide has tripled – in the Netherlands these publications even increased fivefold. Compared to other countries, the Netherlands published a relatively high number of RCTs on GP topics, especially when viewed in relation to gross domestic product and population. The quality of these Dutch studies, measured by the impact factor of the journals in which they were published, corresponded well with the median impact factor of previous studies, with a median annual impact factor ranging from 1.1 to 9.8.(13)

Compared to the research output in other countries, Dutch GP related research is doing well, qualitatively as well as quantitatively.(13) However, the themes of the research corresponded insufficiently with daily practice. One possible explanation could be the guiding influence of research funding, which in principle takes place along four channels. The first source of funding are the own resources of university departments and is limited. The second source of funding are governmental funds, including ZonMw [The Dutch organization for health research and development], and is largely thematically driven by the funding body. For example, programs such as 'Geestkracht' (psychiatric disorders), 'Diabetes Ketenzorg' (type 2 diabetes) and 'Nationaal Programma Ouderenzorg' (care for elderly) have been set up in recent years.(14) The third source of funding are donations to charities like the 'Nederlandse Hartstichting' (Dutch Heart Foundation), the Diabetes Fonds (Diabetes Foundation), the Astma Fonds (Asthma Foundation) and the Hersenstichting (Brain Foundation), and is therefore mainly focused on chronic diseases. The fourth source of funding, commercial research funding, plays only a modest role in GP based research.(5)

Our analysis shows that a considerable amount of research is performed on some ICPC chapters that are not prevalent, whereas some highly prevalent topics are addressed infrequently in research. It may be that a number of questions relevant to GP were studied and answered before 1990, of course. Furthermore, a research agenda is determined by many other factors, such as the relevance of the research question, funding possibilities, the political situation, the chance of publishing research results, and the department's strategy for research. In 2002 ZonMw introduced the so-called 'Programma Alledaagse Ziekten' (programme for Common Diseases, PAZ) to encourage research on common disorders in GP.(14) As a result, skin and subcutaneous tissues (ICPC chapter S), the musculoskeletal system (L), eyes (F), and ears (H) were

studied. The results of the PAZ don't show in our analysis (figure 4) – at least not for the sections S, F and H, perhaps because some studies are not finished yet. Within the PAZ there was relatively little money available for each project and it's possible this could have had its effect on the number and the impact factor of publications.

Limitations

The main limitation of our research is that we did not find all relevant RCTs with the search strategy used. In a random sample that we took by means of an internal check, it turned out that of the 100 RCTs that were published by Dutch GP researchers, only 40 appeared in our database. There are various explanations for this. Firstly, we have not searched in other databases, such as EMBASE and the Cochrane Library, as the funds available did not allow for this. Another explanation lies in the use of PubMed-limits to identify RCTs. These filters select on the basis of tags assigned to a publication, which means that publications that were labelled too late or wrongly could have been missed. We aimed to cover the largest possible area by utilising 'Clinical Trial' as a search term, as well as 'Randomized Controlled Trial'. However, since we haven't found any missing RCTs during our cross-reference check in five meta-analyses, our filter strategy will not have led to major omissions. A third – and in our opinion the most likely – explanation is that it was not apparent from the title or the summary that the character of the research was GP related. This could have resulted in an RCT not being traceable using a search strategy based on domain. Searching on author would have been an alternative, but not all RCTs of GP researchers address a tropic relevant to GP.(6) A fourth explanation is that we have limited ourselves to publications in English and have searched on English synonyms for GP medicine.(9) This excludes foreign-language publications, and also English language publications in countries where general practice is named differently ('community medicine', 'rural health', etc.). There is no optimal search strategy for GP research within online databases. After having collected our data, two articles have been published in which a more comprehensive search strategy was used. In addition to our search strategy, Jelercic (2010) used the terms 'family practice', 'physicians, family', 'rural health care', 'rural health services', 'comprehensive health care', and 'community health services'.(16) Glanville (2011) has developed a highly sensitive search strategy with many separate search terms, applied to six countries with well-developed primary care.(6) However, neither of the search strategies have been validated and their purpose wasn't only to find GP based research, but also research done by GPs. In our view the limitations of our own search strategy are mainly quantitative and systematic. We have no indications that these limitations also have led to a selection in research themes. A third limitation concerns the indicator that we used for research needs from clinical practice, the distribution of clinical topics in GP in the Netherlands: this does not necessarily reflect the necessity for research.(7) Many diseases that are highly prevalent in the field of GP are self-limiting and do not require evidence-based interventions.

Recommendations

Earlier research has shown that the research output of the Netherlands in the field of GP compares very well to that in other countries,(6) but the matching with daily practice could improve.(7) We analysed these developments again, and concluded that the research agenda in GP, despite past initiatives, still not optimally takes daily clinical practice into account.(1,14) In this context, the discontinuation of the PAZ has not been a good development. For general practice it is very important that the research agenda is driven by daily practice, based on questions from daily clinical practice. A database that lists knowledge gaps in the national GP guidelines, the so-called 'lacunebak', is an excellent starting point to identify requirements from daily clinical practice.(15)

REFERENCES

1 Hummers-Pradier E, Beyer M, Chevallier P, Eilat-Tsanani S, Lionis C, Peremans L, et al. The research agenda for general practice/family medicine and primary health care in Europe, Part 1: Background and methodology. Eur J Gen Pract 2009;15:243-50.

2 Van Weel C. The impact of research in primary care and family medicine: The Thomson Reuters Web of Science Subject Category 'Primary Health Care'. Fam Pract 2011;28:239-40.

3 Mant D, Del Mar CB, Glasziou P, Knottnerus A, Wallace P, Van Weel C. The state of primary-care research. Lancet 2004;364:1004-6.

4 Van Weel C, Rosser WW. Improving health care globally: A critical review of the necessity of family medicine research and recommendations to build research capacity. Ann Fam Med 2004;2 Suppl 2:S5-16.

5 Dekker J . Standaarden en onderzoek voor en door huisartsen. Huisarts Wet 2010;53: 47-50.

6 Glanville J, Kendrick T, McNally R, Campbell J, Hobbs FD. Research output on primary care in Australia, Canada, Germany, the Netherlands, the United Kingdom, and the United States: Bibliometric analysis. BMJ 2011;342:d1028.

7 Bakkenist T, Zaat JOM. Wie onderzoekt wat? Een globaal overzicht van huisartsgeneeskundig onderzoek in Nederland. Huisarts Wet 2004;47:555-60.

8 Brozek JL, Akl EA, Allonso-Coello P, Lang D, Jaeschke R, Williams JW, et al. Grading quality of evidence and strength of recommendations in clinical practice guidelines, Part 1 of 3: An overview of the GRADE approach and grading quality of evidence about interventions. Allergy 2009;64:669-77.

9 Ovhed I, Van Royen P, Håkansson A. What is the future of primary care research? Probably fairly bright, if we may believe the historical development. Scand J Prim Health Care 2005;23:248-53.

10 ISI Web of Knowledge [internet]. New York: Thomson Reuters, 2010-2012. http://www.isiknowledge.com, geraadpleegd maart 2010.

11 CIA. The World Factbook [internet]. Washington (DC): Central Intelligence Agency, 2007-2011. https://www.cia.gov/library/publications/the-world-factbook/rankorder/2001rank.html, geraadpleegd april 2011.

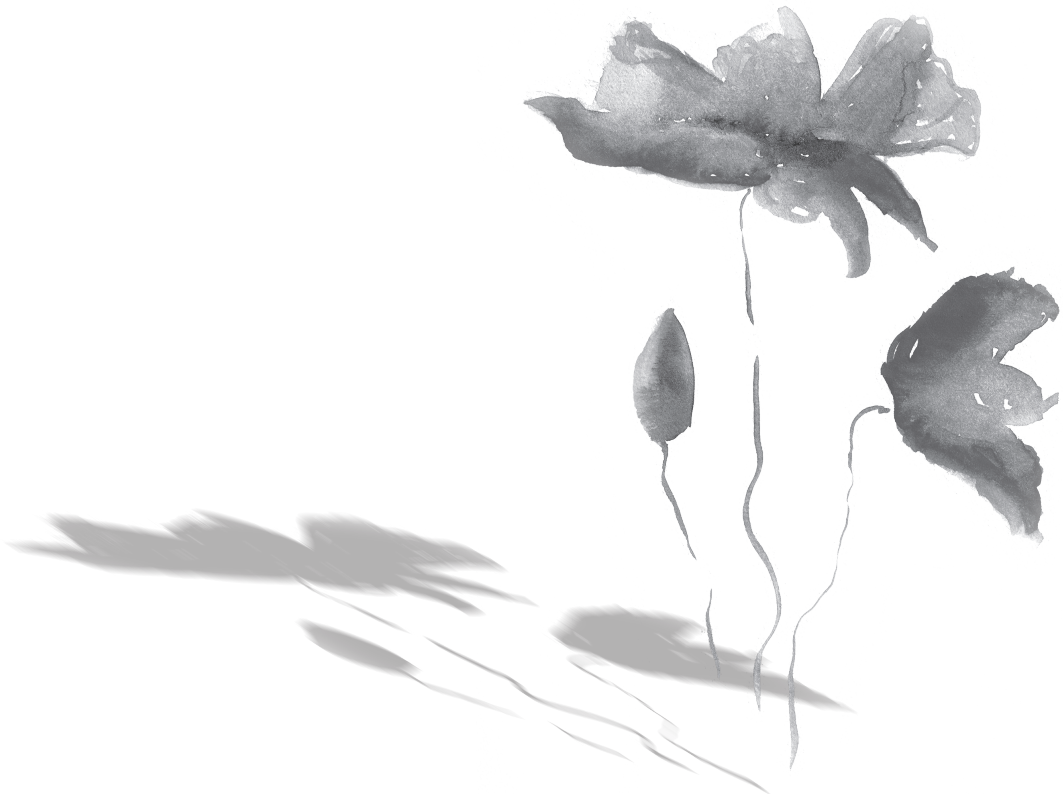
12 International classification of primary care, second edition (ICPC-2). Geneva: World Health Organization, 2003.

13 Van Driel ML, Magin PJ, Del Mar CB, Furler J, De Maeseneer J. Journal impact factor and its importance for AFP. Aust Fam Physician 2008;37:770-3.

14 http://www.zonmw.nl/onderwerpen/cure, geraadpleegd april 2011.

15 Tasche MJA, Oosterberg EH, Kolnaar BGM, Rosmalen CFH. Inventarisatie van lacunes in huisartsgeneeskundige kennis. Zeventig standaarden doorgelicht. Huisarts Wet 2001;44:91-4.

16 Jelercic S, Lingard H, Spiegel W, Pichlhöfer O, Maier M. Assessment of publication output in the field of general practice and family medicine and by general practitioners and general practice institutions. Fam Pract 2010;27:582-9.



CHAPTER 3

Towards efficient use of research resources: a nationwide database of ongoing primary care research projects in the Netherlands

Based on: Kortekaas MF, et al. Towards efficient use of research resources: a nationwide database of ongoing primary care research projects in the Netherlands. Fam Pract 2014;31:229-35.

Kortekaas MF, van de Pol AC,
van der Horst HE, Burgers JS, de Wit NJ

"Doen wat je leuk vindt is vrijheid.
Leuk vinden wat je doet is geluk."
Onbekend

ABSTRACT

Purpose

Although in the last decades primary care research has evolved with great success, there is a growing need to prioritize the topics given the limited resources available. Therefore we constructed a nationwide database of ongoing primary care research projects in the Netherlands, and assessed if the distribution of research topics matched with primary care practice.

Methods

We conducted a survey among the main primary care research centers in the Netherlands, and gathered details of all ongoing primary care research projects. We classified the projects according to research topic, relation to professional guidelines and knowledge deficits, collaborative partners, and funding source. Subsequently, we compared the frequency distribution of clinical topics of research projects to the prevalence of problems in primary care practice.

Results

We identified 296 ongoing primary care research projects from 11 research centers. Most projects were designed as RCT (35%) or observational cohort (34%), and government funded mostly (60%). Thematically most research projects addressed chronic diseases; mainly cardiovascular risk management (8%), depressive disorders (8%), and diabetes mellitus (7%). One fifth of the projects was related to defined knowledge deficits in primary care guidelines. From a clinical primary care perspective, research projects on dermatological problems were significantly underrepresented (p 0.01).

Conclusions

This survey of ongoing projects demonstrates that primary care research has a firm basis in the Netherlands, but has a strong focus on chronic disease. The fit with primary care practice can improve, and future research should address knowledge deficits in professional guidelines more.

INTRODUCTION

Over the past decades, the volume of primary care research output is expanding. A recent bibliometric analysis demonstrates that the Netherlands and the United Kingdom are among the top in primary care research performance, both in quality and in quantity.[1] Practice based research in primary care is essential as it provides the fundamentals for evidence-based practice and can fill in the knowledge deficits in practice guidelines.[2] Currently, resources for clinical research are being cut, and prioritizing research topics is receiving increasing attention.[3] Efficient use of resources for research requires that future projects should address clinically relevant questions, and focus on problems that are frequently encountered in everyday primary care practice.

(Inter) national coordination of the research agenda could encourage efficient use of resources and facilitate research planning in future. For this reason the European General Practice Research Network (EGPRN) proposed to identify research needs by starting with an overview of ongoing research projects per country.[4] Following this suggestion we have constructed a national database of all ongoing general practice based research projects in the Netherlands, that could serve as an example to other countries. We analyzed the characteristics of current research projects, and explored the match between the distribution of research themes and that of clinical topics in everyday primary care practice.

METHODS

Design

Electronic survey among all academic primary care departments and non-academic primary care research centers in the Netherlands, collecting information on all ongoing research projects in primary care.

Definition and selection of primary care research

We defined primary care research as research conducted in primary care patient populations, on a topic relevant for primary care practice, and accounted for externally (either in funding, registration or publication). This included research projects with patients recruited in the primary care setting that address a primary care question and/or projects conducted by a primary care research team. We considered projects to be "ongoing" if patients were currently being enrolled, and/or data were being processed, and funding had ended less than one year before. International projects were also included, provided that patient inclusion took place in the Netherlands as well. We excluded projects that did not meet these criteria.

Data collection

In the Netherlands, primary care research is conducted by eight academic primary care departments, and by a limited number of governmental and private research institutions (Nivel, IQ healthcare and Trimbos institute). We contacted the heads of these 11 research centers, and asked them to participate in this project by providing an overview of their ongoing research projects. Between December 2011 and August 2012, we sent a standardized digital questionnaire to the coordinating researchers of these projects. Non-responders were reminded several times, and if necessary contacted via supervisors. In the survey we gathered characteristics of the researchers (gender, age, function, and professional background) as well as details of the projects (title, start, and (expected) end date). Researchers were asked to classify their projects on seven different aspects: 1) research theme[4] (table 1; ***a.*** clinical research, ***b.*** health services research, e.g. quality of care, communication, and practice organization, ***c.*** research of education and teaching in general practice, ***d.*** combinations of a, b, and c, and ***e.*** other), 2) design (randomized controlled trial (RCT), observational cohort, ecological study, patient-control study, cross-sectional design, other), 3) clinical topic, classified according to the International Classification of Health Problems of Primary Care (ICPC)[5], 4) relation to professional guideline(s) (the 100 existing guidelines of the Dutch College of General Practitioners)[6], 5) matching with existing knowledge deficits in these professional practice guidelines, as defined in the database of the Dutch College of General Practitioners[7], 6) collaboration with different research centers, and 7) the type of funding (institutional, governmental, patient organizations or pharmaceutical company). Researchers could indicate up to three professional guidelines and/or knowledge deficits for each research project. In

Table 1. Definitions of research themes.

Research theme	Definition
1. Clinical research*	1. Research with outcomes at a patient level, measuring patients' health issues including function or quality of life.
2. Health services research*	2. Research focusing on doctor or system related questions and outcomes.
- Quality of care	- Features of care that are important to meet defined or self-evident needs for good care.
- Communication	- Sending and receiving messages.
- Practice organization	- The way a general practice is organized.
3. Research on education and teaching in GP*	3. Research on education and teaching in general practice.
- Education	- Organized, non-incidental communication with the purpose to teach skills, add insight(s), or delegation of knowledge.
4. Combination	4. Combined clinical/ health services / educational research.
5. Other	5. Other research than mentioned above.

* Definitions for general practice/family medicine research, as defined by EGPRN[3]

case of collaborative research projects, projects were assigned to the institution that received the research grant or hosted the project management.

Match between research themes and primary care practice

To analyze the match between the distribution of research themes and that of clinical themes in everyday general practice we compared the frequency (in percentages) of ICPC codes (aggregated by chapter) of the reported research projects with the frequency (in percentages) of reasons for encounter (in ICPC chapters) in primary care practice in the Netherlands.[8] We compared them graphically and assessed the differences using the Chisquared-test.

RESULTS

Researchers

We approached 325 researchers from 11 primary care research centers in the Netherlands, of whom 280 (86%) participated. After exclusion of 31 non-eligible projects, 296 research projects from 256 researchers were included (figure 1). The coordinating researchers were predominantly female (65%), with a mean age of 37.3 years (range 25-65). The professional background of the researchers was mostly medical doctor (55%, of whom 23% GP trainee), health scientist (17%) or epidemiologist (15%). Some combined two (***n***=93), or three (***n***=13) professions. As for their academic position most of the researchers (68%) were PhD students, 14% was post-doc, and 7% was associate or full professor (table 2).

Figure 1. Inclusion of research projects.

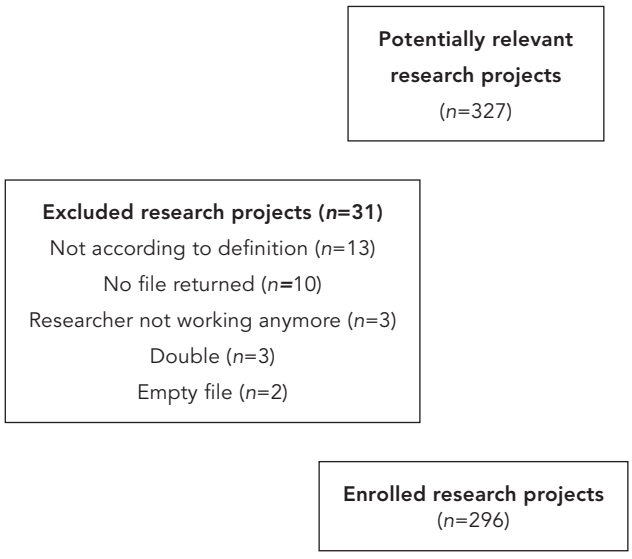


Table 2. Characteristics of researchers.

	N (%)
Gender	
Male	81 (32)
Female	167 (65)
Missing	8 (3)
Age	
Mean	37.3 (25-65)
Missing	11 (4)
Professional background	
- Doctor	137 (55)
- Health scientist	42 (17)
- Epidemiologist	38 (15)
- GP-trainee	31 (12)
- Social scientist	20 (8)
- Psychologist	16 (6)
- Human movement scientist	15 (6)
- Biomedical scientist	12 (5)
- Different	59 (24)
Missing	5 (2)
Academic position	
Number of (combined) positions	
- One	241 (97)
- Two	8 (3)
Type of position	
- Professor	6 (2)
- Associate professor	14 (5)
- Post-doctoral	35 (14)
- PhD student	174 (68)
- Student	2 (1)
- Different	26 (10)
Missing	7 (3)

Research projects

Most projects focused on a clinical topic (48%), followed by those with a theme on health services research (30%), e.g. quality of care (24%), communication (3%), and practice organization (3%). Least research was performed on education and teaching in general practice (1%). (Table 3). Some research projects (12%) covered several themes. Some projects included two (26%) or three (1%) research designs, which resulted in a total number of 373 reported designs for 291 projects. Most projects were designed as RCT (34%) or observational cohort (34%), followed by a patient control (23%), or a cross-sectional study (9%). The majority of research projects (57%) focused on one ICPC code, 10% on two, and 5% on three ICPC-codes. Thirty-six (12%) research projects focused on a topic that could not be captured in a single ICPC-category ("not applicable"), for example a research project about patient safety or mindfulness training for patients with cancer. Most frequently reported clinical topics were cardiovascular conditions (K, 15%), respiratory conditions (R, 14%), psychiatric

conditions (P, 14%), and musculoskeletal conditions (L, 14%). Less frequent topics were hematological conditions (B, 1%) and dermatology (S, 1%). None of the research projects focused on ophthalmology (F) and pregnancy (W).

Relation to professional primary care guidelines and their knowledge deficits

In total, 191 of the 296 research projects (65%) focused on one of the clinical topics of the professional guidelines of the Dutch College of General Practice.[6] Most of these projects (46%) addressed one, 14% addressed two, and 6% three primary care guidelines. The top three of related guidelines were cardiovascular risk management (8%), depressive disorders (8%), and diabetes mellitus (7%), all chronic diseases.[9] Ninety-nine (39%) of the 256 researchers were familiar with the database of existing knowledge deficits in primary care guidelines of the Dutch College. Of the 124 projects under their supervision, only 27 research projects (22%) matched with one or more knowledge deficits from the database.

Collaboration and funding

Most projects (83%) were conducted in collaboration between departments, 49 projects (17%) were single center research. Thirteen projects (4%) reported international collaboration. Most projects (66%) were funded by one funding body, 79 (27%) by two, and 10 (3%) by three different funding bodies. The funding body was mostly the government (57%), followed by the university (38%), and patient organizations (19%). Few projects were funded by the pharmaceutical industry (15%). Twenty-three projects (8%) received international funding.

Match with clinical topics in everyday primary care practice

Figure 2 present the frequency distribution of clinical topics of all research projects identified and the prevalences of reasons for encounter in primary care practice in the Netherlands. We found significant differences ($p < 0.01$); for dermatological conditions (S: 1% in research vs. 15% in clinical practice, p 0.01), the circulatory tract (K: 13% vs. 7%, p 0.03), psychiatric conditions (P: 12% vs. 5%, $p<0.01$), and the endocrine system (T: 8% vs. 3%, $p<0.01$) (web appendix, table 4).

DISCUSSION

This overview of ongoing projects demonstrates that there is a substantial body of research in primary care in the Netherlands, with a wide range of topics. The 11 research institutions involved in primary care research have on average more than 20 ongoing research projects. The majority of these projects have a clinical primary care perspective, a high quality research design (RCT), and is non-commercially funded. However, the present

Table 3. Characteristics of research projects.

	Projects *N* (%)
Theme	
Clinical research	143 (48)
Health services research	89 (30)
- Quality of care	70 (24)
- Communication	9 (3)
- Practice organization	10 (3)
Research on education and teaching in general practice	4 (1)
Combination	34 (11)
Other	24 (8)
Missing	2 (1)
Methodological design	
Number of designs per project	
- One design	212 (72)
- Two designs	76 (26)
- Three designs	3 (1)
Type of designs	
- RCT	102 (34)
- Observational cohort	100 (34)
- Patient control	68 (23)
- Cross-sectional	26 (9)
- Ecological	2 (1)
- Other	75 (25)
Missing	5 (2)
Clinical topics	
Number of ICPC-codes per project	
- None	36 (12)
- One ICPC-code	169 (57)
- Two ICPC-codes	29 (10)
- Three ICPC-codes	16 (5)
Top 3 of clinical topics	
- Cardiovascular (K)	45 (15)
- Respiratory (R)	41 (14)
- Psychiatric (P)/Musculoskeletal (L)	40 (14)
Missing	46 (16)
National guidelines	
Number of national guidelines per project	
- None	53 (18)
- One guideline	137 (46)
- Two guidelines	34 (11)
- Three guidelines	18 (6)
- More than three guidelines	2 (1)
Top 3 of national guidelines	
- Cardiovascular	25 (8)
- Depressive disorders	23 (8)
- Diabetes Mellitus	21 (7)
Missing	52 (18)
Knowledge deficits	
Relation to knowledge deficit	
- No	97 (33)
- Yes	27 (9)
Missing	172 (58)

Table 3. *Continued*

	Projects *N* (%)
Collaboration	
Number of collaborations per project	
- None	49 (17)
- One collaboration	172 (58)
- Two collaborations	60 (0)
Type of collaborations	
Internal	
- Another department (non primary care)	97 (33)
External	
- Another primary care research center	66 (22)
- Another non-primary care research center	103 (35)
Other	26 (9)
Missing	15 (5)
Funding	
Number of funding bodies	
- One funding body	195 (66)
- Two funding bodies	79 (27)
- Three funding bodies	10 (3)
Type of funding	
- Governmental	169 (57)
- Institutional	113 (38)
- Patient organizations	57 (19)
- Pharmaceutical industry	44 (15)
Missing	12 (4)

research agenda shows a strong focus on chronic diseases and a suboptimal match with daily primary care practice and the existing knowledge deficits in clinical guidelines.

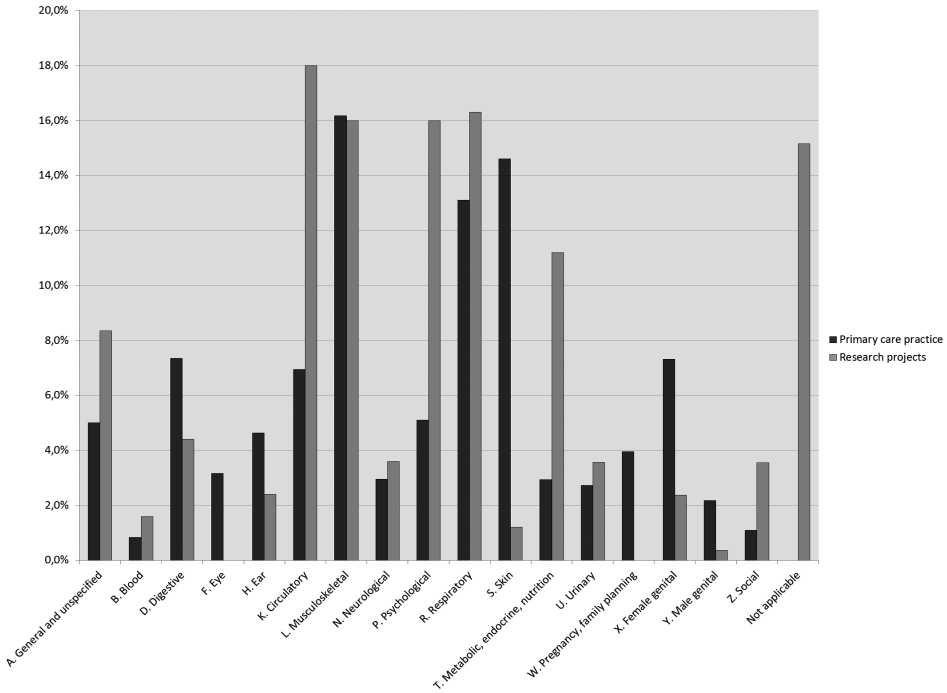
Comparison with existing literature

Our overview is comparable with the database of ongoing clinical research projects in England, as supported by the NIHR Clinical Research Network (CRN).[10] However, projects in this database are neither classified according to ICPC-codes, nor related to (national) guidelines or knowledge deficits.[10] Two earlier reviews analyzed clinical topics classified by ICPC-code, one from UK (1998) and one from Germany (2012).[9, 11] Both reported comparable results regarding the preferred areas of research. Research designs of the studies identified were different: in the UK review only 6% of the projects were RCT, in Germany the majority of the projects had a cross-sectional design.[9, 11] This overview demonstrates that primary care has a firm position in the research field in the Netherlands, with increasing number of RCT's.[12]

Strengths and limitations

This was the first study in the Netherlands providing a nationwide overview of ongoing primary care research. All heads of the research centers cooperated and the response

Figure 2. Clinical topics (ICPC) of research projects compared to the frequency distribution of presented reasons for encounter in primary care practice in the Netherlands[7].



The percentage of research projects sum up above 100 due to research projects focusing on one to three ICPC-codes per project.

rate among the coordinating researchers was high. Our study had a few limitations. First, there was no generally accepted definition for primary care research. Several search strategies have been developed to identify primary care research output, all different, and all with specific limitations.[1; 13-17] We used a broad definition of primary care research to reduce the risk of missing projects. Second, we limited our survey to the primary care research centers in the Netherlands, so we may have missed projects conducted in primary care by other professionals and projects coordinated from hospitals. However, assuming that in most research projects in primary care, primary care researchers will be involved, we expect the number of missing projects to be limited. Third, we used the frequency distribution of presented reasons for encounter in primary care practice as a proxy indicator for the spectrum of primary care research required, whereas the true needs depend on other factors as well, such as knowledge deficits, disease burden, and healthcare costs.[18] Finally, our attempt to match ongoing projects with the database of knowledge deficits in the professional guidelines of the Dutch college was not very successful, as most researchers were not familiar with this, and as the quality of the database was suboptimal.

Table 4. Clinical topics (ICPC) of research projects compared to the frequency distribution of presented reasons for encounter in primary care practice in the Netherlands.[7]

ICPC code	Primary care practice	Research projects	Difference	p value*
A. General and unspecified	5%	8%	3%	1.0
B. Blood	1%	1%	1%	1.0
D. Digestive	7%	4%	-3%	1.0
F. Eye	3%	0%	-3%	1.0
H. Ear	5%	2%	-2%	1.0
K. Circulatory	7%	16%	9%	0.03
L. Musculoskeletal	16%	15%	-2%	1.0
N. Neurological	3%	3%	0%	1.0
P. Psychological	5%	15%	10%	< 0.01
R. Respiratory	13%	15%	2%	1.0
S. Skin	15%	1%	-14%	0.01
T. Metabolic, endocrine, nutrition	3%	10%	7%	< 0.01
U. Urinary	3%	3%	1%	1.0
W. Pregnancy, family planning	4%	0%	-4%	1.0
X. Female genital	7%	2%	-5%	1.0
Y. Male genital	2%	0%	-2%	1.0
Z. Social	1%	3%	2%	1.0

* Corrected for multiple comparisons with Bonferoni method

Implications for the future

In our view a database of ongoing primary care research is useful, as it can help to monitor ongoing research projects, asses if actual research needs in primary care are being met, and guide researchers, health policy makers, and (governmental) funding agencies in future research planning. Our database could serve as an example for other countries, and could facilitate (inter)national collaboration in primary care research, as advocated recently.[19] Some conditions are crucial for successful use, such as keeping the database up to date, which needs commitment and continuous input from all researchers involved.

As funding sources for primary care research are being reduced[3], prioritization of themes and topics is of paramount importance. To enable a better match between future primary care research and the needs from primary care practice, funding and project planning could be linked to existing knowledge deficits in clinical practice guidelines. This requires a dynamic database, not only of ongoing research projects, but also of knowledge deficits, derived from existing clinical practice guidelines. Increasing awareness among researchers, health policy makers, and funding organizations could enhance efficient use of resources for research projects meeting the needs in primary care practice better.

REFERENCES

1 Glanville J, Kendrick T, McNally R, Campbell J, Hobbs FD. Research output on primary care in Australia, Canada, Germany, the Netherlands, the United Kingdom, and the United States: bibliometric analysis. BMJ 2011;342:d1028.

2 Sackett DL. Evidence-based medicine. Semin Perinatol 1997 Feb;21(1):3-5.

3 Funding of US biomedical research, 2003-2008. Dorsey ER, de Roulet J, Thompson JP, Reminick JI, Thai A, White-Stellato Z, Beck CA, George BP, Moses H 3rd. JAMA. 2010 Jan 13;303(2):137-43.

4 Hummers-Pradier E, Beyer M, Chevallier P, Cos X, Eilat-Tsanani S, Waltraud F, et al. Research Agenda for General Practice / Family Medicine and Primary Health Care in Europe, European General Practice Research Network EGPRN, Maastricht/Hannover, 2010.

5 International classification of health problems of primary care (ICPC), second edition (ICPC-2). Geneva: World Health Organisation, 2003.

6 http://nhg. artsennet. nl/kenniscentrum/k_richtlijnen/k_nhgstandaarden. htm.

7 http://nhg. artsennet. nl/kenniscentrum/k_richtlijnen/k_nhgstandaarden/k_lacunes. htm.

8 JCC Braspenning, FG Schellevis, RPTM Grol (redactie). Tweede Nationale Studie naar ziekten en verrichtingen in de huisartspraktijk. Kwaliteit huisartsenzorg belicht. Utrecht/ Nijmegen:NIVEL/WOK, 2004.

9 Schneider A, Grossmann N, Linde K. The development of general practice as an academic discipline in Germany - an analysis of research output between 2000 and 2010. BMC Fam Pract 2012;13(1):58.

10 http://www nihr ac uk/Pages/NRRArchive aspx.

11 Thomas T, Fahey T, Somerset M. The content and methodology of research papers published in three United Kingdom primary care journals. Br J Gen Pract 1998 May;48(430):1229-32.

12 Kortekaas MF, Meijer A, Van de Pol AC, De Wit NJ. RCT's in de huisartsgeneeskunde, 1990-2010. Huisarts Wet 2012;55(11):486-91.

13 Jelercic S, Lingard H, Spiegel W, Pichlhofer O, Maier M. Assessment of publication output in the field of general practice and family medicine and by general practitioners and general practice institutions. Fam Pract 2010 Oct;27(5):582-9.

14 Ovhed I, van RP, Hakansson A. What is the future of primary care research? Probably fairly bright, if we may believe the historical development. Scand J Prim Health Care 2005 Dec;23(4):248-53.

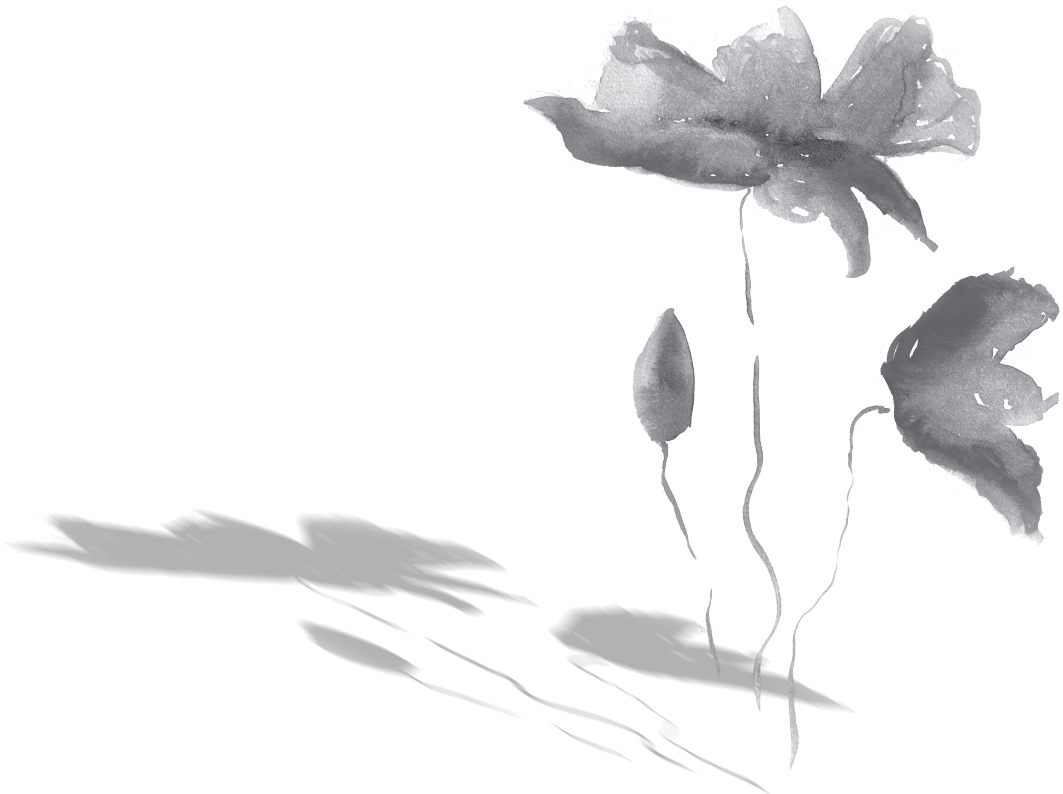
15 Askew DA, Glasziou PP, Del Mar CB. Research output of Australian general practice: a comparison with medicine, surgery and public health. Med J Aust 2001 Jul 16;175(2):77-80.

16 Merenstein J, Rao G, D'Amico F. Clinical research in family medicine: quantity and quality of published articles. Fam Med 2003 Apr;35(4):284-8.

17 Rosser WW, Starkey C, Shaughnessy R. The pond is wider than you think! Problems encountered when searching family practice literature. Can Fam Physician 2000 Jan;46:103-8.

18 Pinnock H, Ostrem A, Rodriguez MR, Ryan D, Stallberg B, Thomas M, et al. Prioritising the respiratory research needs of primary care: the International Primary Care Respiratory Group (IPCRG) e-Delphi exercise. Prim Care Respir J 2012 Mar;21(1):19-27.

19 Research funding. Toward variable funding for international science. Edler J. Science. 2012 Oct 19;338(6105):331-2. doi: 10.1126/science.1221970.



CHAPTER 4

Clinical queries and information seeking in clinical practice by trainees in a hospital-based specialty and trainees general practitioner: A systematic review

Submitted for publication: Clinical queries and information seeking in clinical practice by trainees in a hospital-based specialty and trainees general practitioner: A systematic review.

Kortekaas MF, Hooft L, Berkers N,
Hoes AW, de Wit NJ, Bartelink MEL

"Als je de waarheid hebt gevonden,
heb je niet goed genoeg gezocht."
Socrates

ABSTRACT

Background and aim

In clinical practice physicians frequently pursue answers to clinical questions, e.g. on the value of diagnostic tests or effect of therapies, to optimize patient care. Trainees, at the beginning of their career, may experience different information needs and use different information sources. Trainees general practitioner (GP) may differ from trainees in a hospital-based specialty because of the different setting and wide variation in patient consultations. We conducted a systematic review to assess present knowledge about the frequency of clinical queries encountered in daily practice and information seeking behaviour by trainees GP and trainees in a hospital-based specialty.

Methods

We searched PubMed and Embase in November 2014 and included studies published after 1996 that described clinical question frequencies and information seeking behaviour of trainees. Two researchers independently selected studies and extracted frequencies and percentages on the (relative) number of clinical queries, searches performed, and answers retrieved. We further looked into types of queries, resources used, moments of searching, reasons to search or refrain from searching, time spent searching, and the impact of the retrieved answer.

Results

We identified 9,668 studies, of which 35 met the eligibility criteria. Eleven studies included trainees GP. Instruments most frequently used were questionnaires. Ten studies assessed information needs and seeking behaviour throughout the process of patient management, with a median of 403 included patients (IQR 234-13043). On average, trainees had one clinical query per patient (N=3 studies; 95%CI 0.48-1.52). Most frequently used information sources were senior colleagues (25-66%). Information sources were used 0.83 times per query (N=6 studies; 95%CI 0.57-1.09) and answers retrieved for 0.79 of the searches (N=5 studies; 95%CI 0.76-0.82). Searches were usually performed (5-15 minutes) during patient consultations. Primary reason to refrain from searching was lack of time. There were no clear differences between studies performed in trainees GP and trainees in a hospital-based specialty, though trainees GP tended to have more clinical queries and performed more searches.

Conclusion

Although the studies about trainees' information needs and seeking behaviour differed greatly in design, setting, participants, sample size, instruments, and outcome measures, both trainees GP and trainees in a hospital-based specialty frequently have queries in daily clinical practice, and use information sources and retrieve answers for the vast majority of these encountered queries. However, the use of evidence-based resources could be improved.

INTRODUCTION

For many clinical questions practising physicians have in daily practice, e.g. on the value of a diagnostic test or about the effect of a specific therapy, the answers are not readily available. A recent review by Del Fiol et al (2014) showed that doctors have on average 0.6 clinical questions per patient consultation, that they pursue answers for half of these questions, and that they find answers to 80% of these questions.(1) To seek answers to these queries doctors primarily use text books or human resources.(1,2) The introduction of evidence-based medicine (EBM) and the increasing availability of online information resources seem not to have changed the number of searches, or the type of information resources used.(1–3) In fact, the broad access to online information resources may have made it harder to actually identify the required information, as many doctors may feel flooded by the abundance of online information.(2)

Learning to address clinical questions is essential in medical education. Trainees learn to work according to the principles of EBM during their competency-based postgraduate training programmes.(4) An essential skill in this programme is to search for evidence to answer clinical queries encountered in daily clinical practice, facilitating so called "just-in-time learning".(5) Understanding the information needs and seeking behaviour of trainees is instrumental to further optimize EBM training.(6) Trainees probably differ from experienced doctors in their information needs and seeking behaviour, as the first have limited clinical experience, they are better trained in practicing EBM, and they are accustomed to using internet based resources.(2) In addition, trainees general practice (GP) may differ from trainees in a hospital-based specialty. Learning in a hospital takes place in a complex clinical environment with multiple supervisors, while GP training is focused on the interaction between one trainee and one supervisor, and deals with a broader spectrum of reasons for encounter.(7)

We conducted a systematic review on the frequency of clinical questions encountered in daily practice and subsequently information seeking behaviour of trainees GP and trainees in a hospital-based specialty. We focussed on the frequency of clinical questions and seeking behaviour throughout the process of patient management as primary outcome. Additionally we collected data on the types and topics of clinical queries, the sources used to obtain information, the moment searches were performed, the reasons to search or to refrain from searching, the duration of a search, and the (perceived) impact of retrieved answers.

METHODS

Search strategy and study selection

We conducted a systematic search in PubMed and Embase on November 5th 2014 (January 1996 through November 2014), with no language restrictions. We started our search from 1996, as we considered the impact of the world wide web on seeking

behaviour to be established from 1996 onwards.(2) Two investigators (MK and NB) independently abstracted the articles that met the inclusion criteria. Disagreement was resolved by seeking consensus or, in case of persistence, a third investigator (MB) made a final decision. We attempted to contact the corresponding author when articles could not be obtained through institutional holdings available to us. We collected and organized important information and developed a picture of the existing evidence base and gaps, thereby using a broad research question. We included observational and intervention studies that described frequencies of clinical queries and information seeking behaviour of physicians in postgraduate training (trainees), including interns, residents, registrars, and junior doctors. Seeking behaviour was defined as "questions about medical knowledge that could potentially be answered by general sources such as textbooks and journals, not questions about patient data that would be answered by the medical record."(8) Examples of clinical queries are shown in figure 1. Seeking behaviour had to be reported at one or more of three possible aspects of the search process; clinical queries, searches performed, or answers retrieved. The exact searching strategy is shown in table 1.

Data abstraction

Study characteristics and outcomes of interest were extracted using standardized abstract forms by two independent researchers (MK and NB). Extracted characteristics were author, year of publication, year of data collection, study location, study design, study population (number, medical specialty, level of degree), aims(s) of the study, outcomes measures, and method(s) used to assess outcomes with. When applicable we also extracted data of interventions, including duration, frequency, and follow-up. Our primary outcomes of interest were frequencies of clinical queries encountered in daily practice and subsequent information seeking behaviour throughout the process

Figure 1. Seeking behaviour: Examples of different types of questions.

Diagnostic question/ What is the added value of assessing NT-proBNP levels to exclude or detect heart failure in patients with shortness of breath on exertion?

Etiologic question/ Is it possible that a measles infection causes a higher incidence of inflammatory bowel disease in children?

Therapeutic question/ What is the effect of antibiotics compared "watchful waiting" on the total duration of illness in children with acute otitis media?

Prognostic question/ What is the predictive value of the duration and height of fever in children with a history of bacterial meningitis on their school results?

of patient management. We only used baseline data that were collected before an educational intervention, as this may have altered results. If more participants besides trainees were analysed in studies, reported outcomes had to be retraceable to trainees. We extracted information on 1) the number of patients, 2) the number of clinical queries per patient, 3) the number of searches performed per query, and the 4) the number of retrieved answers per search. Secondary outcomes included 1) the types of clinical queries (e.g. diagnostic, etiologic, prognostic, therapeutic, harm, administrative, prevention, other), 2) the topics of clinical queries (e.g. clinical fields such as cardiology, urinary tract, gynaecology, etc.), 3) the sources used to obtain information (e.g. text books, colleagues, internet, PubMed, etc.), 4) the moment searches were performed (e.g. before, during or after patient consultations, later on the day, at home, etc.), 5) reasons to search (to check, to expand knowledge, etc.), 6) reasons to refrain from searching (e.g. little relevance of the question, lack of time, etc.), 7) the duration of a search (in minutes), and 8) the impact of retrieved answers (e.g. improvement of the clinical decision, confirmation, expanding knowledge, reassurance, etc.).

Data analysis

Because of the large variation in observed trainees, study designs, and outcome measurements no attempt was made to perform a meta-analysis of methodological features and contextual factors associated with the frequency of questions. We categorized data according to the participants studied (trainees GP or trainees in a hospital-based specialty), and the methods used to obtain data. For the latter we used the five categories as defined by Del Fiol et al: 1) interviews, 2) self-reports (i.e. logbooks), 3) direct observations, 4) analysis of inquiries submitted to information services (information services) and 5) analysis of online information source use logs (search logs).(1) As we also included studies that did not obtain data throughout the process of patient management, we added two other categories; 6) assignments or exercises (test), and 7) questionnaires. Data on primary outcomes were reported in proportions with means and 95% confidence intervals (95%CI), data on secondary outcomes in percentages.

RESULTS

Selected studies

A total of 15,155 potentially relevant studies were retrieved, of which 9,668 were unique (see flowchart, figure 2). Cross-reference checking revealed three additional studies.(9–11) One study was added on expert's advice.(12) Based on title and abstract, 182 studies were included for full-text screening. Five studies could not be procured and were thus not included in the review.(13–17) Fifty-three studies were then assessed for eligibility. Thirty-five studies provided relevant data on one or more of the outcomes of interest for trainees and were included in the current evaluation.

Table 1. Search terms for MEDLINE and EMBASE.

Database	Search	Hits (dated 05-11-2014)
MEDLINE (PubMed)	((((((information[Title/Abstract] OR evidence[Title/Abstract] OR literature[Title/Abstract])) **AND** (need[Title/Abstract] OR needs[Title/Abstract] OR retrieval[Title/Abstract] OR retrieve[Title/Abstract] OR retrieving[Title/Abstract] OR search[Title/Abstract] OR searches[Title/Abstract] OR searching[Title/Abstract] OR read[Title/Abstract] OR reading[Title/Abstract] OR "conduct search"[Title/Abstract] OR "conduct searches"[Title/Abstract] OR "conducting search"[Title/Abstract] OR "conducting searches"[Title/Abstract])) OR (((clinical[Title/Abstract] OR medical[Title/Abstract])) **AND** (question[Title/Abstract] OR questions[Title/Abstract] OR query[Title/Abstract] OR queries[Title/Abstract] OR problem[Title/Abstract] OR problems[Title/Abstract])) OR (((medical[Title/Abstract] OR evidence[Title/Abstract] OR EBM[Title/Abstract] OR EBP[Title/Abstract])) **AND** (database[Title/Abstract] OR databases[Title/Abstract])) OR (((EBM[Title/Abstract] OR "evidence based medicine"[Title/Abstract] OR EBP[Title/Abstract] OR "evidence based practice"[Title/Abstract] OR search[Title/Abstract] OR searching[Title/Abstract] OR seek[Title/Abstract] OR seeking[Title/Abstract] OR information[Title/Abstract] OR questioning[Title/Abstract])) **AND** (Behavior[Title/Abstract] OR behaviour[Title/Abstract])) OR (((("evidence based"[Title/Abstract] OR information[Title/Abstract] OR online[Title/Abstract] OR EBM[Title/Abstract] OR EBP[Title/Abstract] OR knowledge[Title/Abstract])) **AND** (Resource[Title/Abstract] OR source[Title/Abstract] OR resources[Title/Abstract] OR sources[Title/Abstract]))))) **AND** ((resident[Title/Abstract] OR residents[Title/Abstract] OR preceptor[Title/Abstract] OR preceptors[Title/Abstract] OR "family medicine teaching clinics"[Title/Abstract] OR "family medicine teaching units"[Title/Abstract] OR "Primary Care Teaching Units"[Title/Abstract] OR physician[Title/Abstract] OR physicians[Title/Abstract] OR residency[Title/Abstract] OR residencies[Title/Abstract] OR internship[Title/Abstract] OR internships[Title/Abstract] OR clerkship[Title/Abstract] OR clerkships[Title/Abstract] OR clinician[Title/Abstract] OR clinicians[Title/Abstract] OR doctor[Title/Abstract] OR doctors[Title/Abstract] OR trainee[Title/Abstract] OR trainees[Title/Abstract] OR "junior medical staff"[Title/Abstract] OR "junior medical officer"[Title/Abstract] OR "junior medical officers"[Title/Abstract] OR registrar[Title/Abstract] OR registrars[Title/Abstract] OR "postgraduate medical student"[Title/Abstract] OR "postgraduate medical students"[Title/Abstract] OR "general practice"[Title/Abstract] OR "general practicioner"[Title/Abstract] OR "general practicioners"[Title/Abstract] OR "general practitioner"[Title/Abstract] OR "general practitioners"[Title/Abstract] OR "family medicine"[Title/Abstract] OR "family medicines"[Title/Abstract] OR "family practice"[Title/Abstract] OR "family practices"[Title/Abstract]))) **AND** (EBM[Title/Abstract] OR EBP[Title/Abstract] OR "evidence based"[Title/Abstract] OR EBCP[Title/Abstract] OR EBHC[Title/Abstract]) **AND** ("1996/01/01"[Date - Publication] : "3000"[Date - Publication])	6375

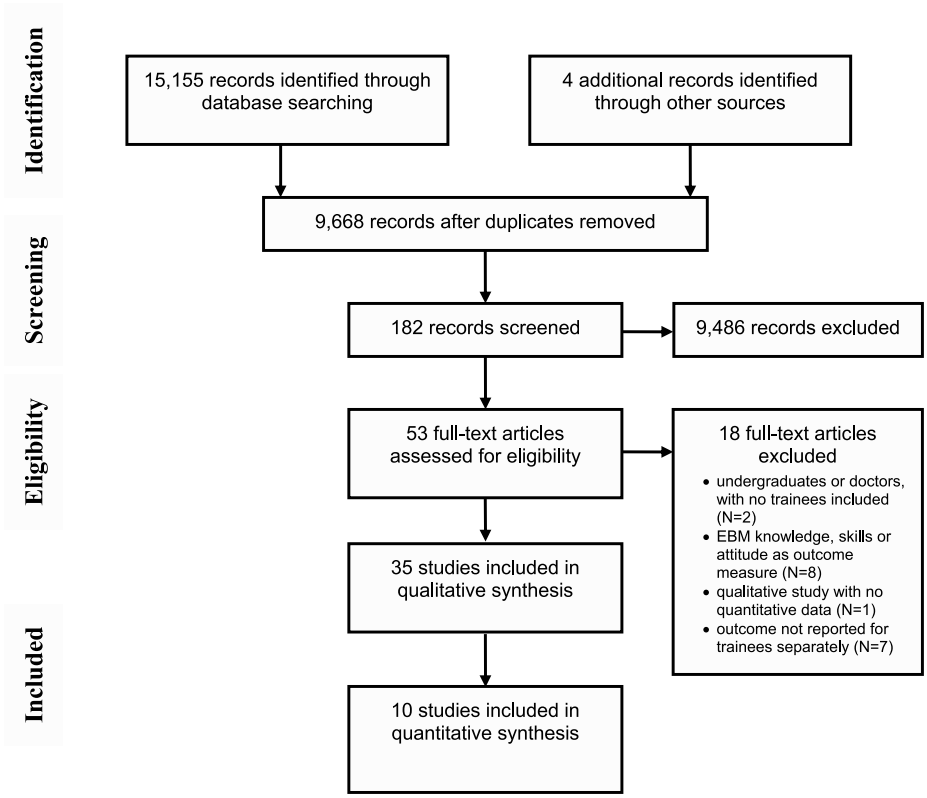
Table 1. *Continued*

Database	Search	Hits (dated 05-11-2014)
EMBASE (Ovid)	information:ab,ti OR evidence:ab,ti OR literature:ab,ti **AND** (need:ab,ti OR needs:ab,ti OR retrieval:ab,ti OR retrieve:ab,ti OR retrieving:ab,ti OR search:ab,ti OR searches:ab,ti OR searching:ab,ti OR read:ab,ti OR reading:ab,ti OR 'conduct search':ab,ti OR 'conduct searches':ab,ti OR 'conducting search':ab,ti OR 'conducting searches':ab,ti) OR (clinical:ab,ti OR medical:ab,ti **AND** (question:ab,ti OR questions:ab,ti OR query:ab,ti OR queries:ab,ti OR problem:ab,ti OR problems:ab,ti)) OR (medical:ab,ti OR evidence:ab,ti OR ebm:ab,ti OR ebp:ab,ti **AND** (database:ab,ti OR databases:ab,ti)) OR (ebm:ab,ti OR 'evidence based medicine':ab,ti OR ebp:ab,ti OR 'evidence based practice':ab,ti OR search:ab,ti OR searching:ab,ti OR seek:ab,ti OR seeking:ab,ti OR information:ab,ti OR questioning:ab,ti **AND** (behavior:ab,ti OR behaviour:ab,ti)) OR ('evidence based':ab,ti OR information:ab,ti OR online:ab,ti OR ebm:ab,ti OR ebp:ab,ti OR knowledge:ab,ti **AND** (resource:ab,ti OR source:ab,ti OR resources:ab,ti OR sources:ab,ti)) **AND** (resident:ab,ti OR residents:ab,ti OR preceptor:ab,ti OR preceptors:ab,ti OR 'family medicine teaching clinics':ab,ti OR 'family medicine teaching units':ab,ti OR 'primary care teaching units':ab,ti OR physician:ab,ti OR physicians:ab,ti OR residency:ab,ti OR residencies:ab,ti OR internship:ab,ti OR internships:ab,ti OR clerkship:ab,ti OR clerkships:ab,ti OR clinician:ab,ti OR clinicians:ab,ti OR doctor:ab,ti OR doctors:ab,ti OR trainee:ab,ti OR trainees:ab,ti OR 'junior medical staff':ab,ti OR 'junior medical officer':ab,ti OR 'junior medical officers':ab,ti OR registrar:ab,ti OR registrars:ab,ti OR 'postgraduate medical student':ab,ti OR 'postgraduate medical students':ab,ti OR 'general practice':ab,ti OR 'general practicioner':ab,ti OR 'general practicioners':ab,ti OR 'general practitioner':ab,ti OR 'general practitioners':ab,ti OR 'family medicine':ab,ti OR 'family medicines':ab,ti OR 'family practice':ab,ti OR 'family practices':ab,ti) **AND** (ebm:ab,ti OR ebp:ab,ti OR 'evidence based':ab,ti OR ebcp:ab,ti OR ebhc:ab,ti) **AND** [1996-2014]/py	8780
TOTAL		15155

Study characteristics

Main characteristics of included studies are listed in 2. The 35 included studies were published between 1999 and 2014 and reported on data collected between 1994 and 2012. Studies were geographically from 10 different countries. About half of the studies were conducted in the United States of America, where the rest were mostly from Europe. Of the 35 included studies, 9 reported data on trainees GP (12,18–24), 24 on trainees in a hospital-based specialty, and two on both (25,26). Sample sizes of participants differed considerably (range 23-610). Instruments used to assess outcomes in the different studies also varied; four used interviews (9,12,25,27), one self-reports (25), six direct observations (12,18,22,23,27,28), three information services, five search logs (10,29–32), three educational assignments to search for evidence to answer specific questions (21,33,34), and twenty administered questionnaires to the trainees(11,20,22–24,26,32,35–47). In

Figure 2. Flowchart showing authors' searches and selection of studies for inclusion.



seven of the studies more than one method was used.(12,22,23,25,27,32,35) There was a great diversity amongst the reported outcomes. Different outcomes were reported, but also different methods were used for measuring the same outcome.

Information seeking behaviour throughout the process of patient management

Twenty-two studies reported data on seeking behaviour throughout the process of patient management, of which six included trainees GP (18,19,22–24,48), 15 trainees in a hospital-based specialty (9–11,27,29,30,32,35–39,43,46,49), and one included both (25). Ten of these studies reported data on one or more of the primary outcomes; three studies on trainees GP, and seven on trainees in a hospital-based specialty. (9,11,18,22,23,27,30,34,35,49) Data on these primary outcomes are shown in table 3.

Trainees GP

The two studies reporting on the number of clinical queries reported that trainees GP had an average of 0.77 and 1.53 clinical queries per patient (mean of the two studies

1.15, 95%CI 0.41-1.89), and performed on average 0.77 and 1.25 searches per clinical query (mean 1.01, 95%CI 0.54-1.48).(18,23) Zwolsman et al. observed that trainees GP searched less than 0.03 times per patient, but did not report on the number of searches related to queries.(12) Allan et al. showed that trainees GP retrieved on average 0.83 answers to searches performed.(18) McCord et al. observed an average of 1.6 retrieved answers per patient, but did not report on the number of queries or searches performed.(22)

Trainees in a hospital-based specialty

For trainees in a hospital-based specialty there was only one study that reported on the number of clinical queries per patient consultation, reporting an average of 0.69 queries (95%CI 0.61-0.78) per patient, based on 403 patient consultations.(27) The number of performed searches was reported by four studies and ranged from 0.29 to 0.97 searches per clinical query (mean 0.74, 95%CI 0.43-1.05).(9,11,27,49) Methods used to obtain data differed, with lowest numbers of searches measured during observations (27), and highest when data were obtained through a questionnaire (11). Chisolm et al. reported that trainees in a hospital-based specialty searched 1.4 times per patient, as measured by a search log.(30) The average proportion of searches in which an answer was retrieved in studies (N=4) ranged from 0.76-0.83 (mean 0.79, 95%CI 0.76-0.82) obtained in information services, interviews, or a questionnaire. (9,11,35,49) Although Crowley et al. reported on data obtained during the introduction of a web-based system collecting clinical queries, which could be regarded as an intervention, results did not differ from the other studies.

Secondary outcomes of interest

Types and topics of clinical queries

Eight studies reported on the types of clinical queries, of which three included trainees GP (table 3).(11,18,19,21,27,33,34,49) Most queries were therapeutic (N=8 studies, 36-59%), closely followed by diagnostic (N=7 studies, 15-43%), with no large differences between trainees GP and trainees in a hospital-based specialty, or among instruments used to obtain data. Most queries involved cardiology (N=5 studies, 9-15%), gastroenterology (N=4 studies, 4-20%), gynaecology (N=3 studies, 2-12%), or neurology (N=3 studies, 5-27%).

Information sources used

Thirty-one studies (of which 8 were GP studies) reported information on the information sources used. Most frequently used information sources were colleagues, with percentages varying between 44 and 66% for trainees GP (18,22,23), and between 25 and 58% for trainees in a hospital-based specialty (9,11,27,39,46). Risahmawati et al. reported that when trainees have a question 50% of them always consults a senior doctor, whereas 32% of the trainees always consult a colleague.(41) Trainees

Table 2. Studies included in review.

Author	Journal	Year of publication	Study location	Study design
Allan GM et al.	Canadian Family Physician, web exclusive	2012	Canada	Prospectice cohort study
Bergus GR et al.	Familiy Medicine	2005	USA	Prospectice cohort study
Brennan N et al	Health Information and Libraries Jourl	2014	UK	Cross-sectional study
Cabell CH et al.	J Gen Intern Med	2001	USA	RCT
Chisholm R et al.	AMIA Annu Symp Proc	2012	USA	Cross-sectional study
Coulthard MG et al.	Evidence-Based Medicine	2001	Australia	Prospectice cohort study
Crowley SD et al.	Academic Medicine	2003	USA	Prospectice cohort study
d'Allesandro MP et al.	Acad Radiol	1999	USA	Cross-sectional study
Duran-Nelson A et al.	Academic Medicine	2013	USA	Cross-sectional study
Feldstein DA, et al.	BMC Medical Education	2010	USA	Prospectice cohort study
Forrest M et al.	Health Libraries Review	2000	UK	Cross-sectional study
Grad R et al	Family Medicine	2001	Canada	Prospectice cohort study
Green ML et al.	The American Journal of Medicine	2000	USA	Prospectice cohort study
Hadley JA et al.	BMC Medical Education	2007	UK	Cross-sectional study
Hayward RS et al.	AMIA Annu Symp Proc	2006	Canada	Prospectice cohort study
Jeve YB et al.	JRSM Short Reports	2013	UK	Cross-sectional study
Kim S et al.	J Gen Intern Med	2008	USA	RCT
Korownyk C et al.	Medical Teacher	2013	Canada	Cross-sectional study
Leon SA et al.	BMC Medical Informatics and Decision Making	2007	USA	Prospectice cohort study

Study population	Aim(s) of the study and outcome measures	Methods used to assess outcomes with
GP trainees	Use of resources, before and after EBM workshop (and between residents in semi-independent clinics (IMGs) and under normal supervision)	Observational
GP trainees	Quality of clinical questions asked (PICO)	Information Services
Qualified doctors across all grades and medical students, GP trainees and trainees from other specialties	Use of resources for accessing information; what, and why	Self-report & interview
Internal Medicine residents	Literature search activity	Search log
All clinicians (students, residents, staff) working or consulted at ER	Search terms and website usage	Search log
Paediatric registrars and consultants	Seeking behaviour (frequency of questions, moment of searching, resources used, access to databases, questions on the evidence-based information in four clinical scenarios)	Questionnaire & Information Services
Internal Medicine residents	Clinical questions	Information Services
Radiology trainees	Seeking behaviour	Interview
Internal Medicine residents	Resources used for Point-of-Care (POC) decision making	Questionnaire
Internal Medicine residents	Self-reported literature searching and resource use	Questionnaire
Trainees GP and trainees from other specialties	Information needs of doctors-in-training and identification of their preferred sources of information	Questionnaire
Trainees GP	1) Six domains of skills (pre and post-course), and 2) self-reported use of resources, up to 2 years post-course	Questionnaire
Internal Medicine residents	Seeking behaviour in clincial practice	Observational & interview
Junior doctors from various specialties	Assessment of literature search behaviour	Questionnaire
Internal Medicine physisicians and trainees	The use of clinical information systems, decision-support tools and evidence resources	Search log
Gynaecology trainees	Skills, knowledge and attitude regarding different components of EBM and EBM-teaching	Questionnaire
Internal Medicine residents	Primary: Use of evidence-based resources. Secondary: scores on vignettes and EBBM-knowledge	Search log
GP trainees	Descriptive analysis of BEARS	Assignment
Internal Medicine residents (in 290-bed acute care teaching hospital and regiol referral center)	Point-of-Care wireless Internet Access using smart phones for information retrieval during daily clinical rounds and academic activities	Questionnaire & Search log

Table 2. *Continued*

Author	Journal	Year of publication	Study location	Study design
Martinez-Silveira MS et al.	J Med Libr Assoc	2005	Brazil	Cross-sectional study
McCord G et al	Academic Medicine	2007	USA	Prospectice cohort study
Phua J et al.	Medical Teacher	2007	Singapore	Cross-sectional study
Ramos K et al	Family Medicine	2003	USA	Cross-sectional study
Risahmawati Rm et al.	BMC Research Notes	2011	Japan	Cross-sectional study
Sadeghi-Ghyassi F et al.	Journal of Evaluation in Clinical Practice	2012	Iran	Cross-sectional study
Schilling LM et al.	Academic Medicine	2005	USA	Cross-sectional study
Shirkhedkar P & Day AS	Informa Healthcare	2008	Australia	Cross-sectional study
Slawson DC et al.	J Am Board Fam Pract	1999	USA	Prospectice cohort study
Stark R et al.	J Gen Intern Med	2007	USA	RCT
Thom DH et al.	BMC Medical Education	2004	USA	Prospectice cohort study
Tilburt JC et al.	Journal of Evaluation in Clinical Practice	2007	USA	Qualitative study design
Veness M et al.	Radiation Oncology	2003	Australia and New Zealand	Cross-sectional study
Westbrook JI et al.	Internal Medicine Jourl	2005	Australia	Cross-sectional study
Yousefi-Nooraie R et al.	Journal of Evaluation in Clinical Practice	2007	Iran	Cross-sectional study
Zwolsman SE et al.	Perspect Med Educ	2013	The Netherlands	Qualitative study design

Study population	Aim(s) of the study and outcome measures	Methods used to assess outcomes with
Residents from several specialties	Information needs and behaviour	Questionnaire
GP trainees	Type of information sources used by EBM-trained trainees GP at the point of care: resources, time, retrieval location, number of patients	Observational & Questionnaire
Residents from different specialties (in a tertiary hospital)	Use of medical-information resources	Questionnaire
GP trainees & GPs	Number of clinical questions, sources consulted, search times, satisfaction with answers	Observational & Questionnaire
Residents from several specialties	EBM attitude and knowledge and barriers to use EBM	Questionnaire
Faculty staff and residents from several specialties	Information-seeking behaviour	Questionnaire
Internal Medicine	How does answering a patient-specific clinical question affect residents' patient care decisions	Assignment
Paediatric Junior Medical Officers (JMO) or JMO from other specialties when they were currently posted at paediatrics as part of a paediatric rotation (orthopedics, surgery, ophthalmology, psychiatry, GP)	Use of online information retrieval systems	Questionnaire
GP trainees and GPs	General attitude (7 items), ability to evaluate clinical research (5 items), use (12 items)	Questionnaire
Internal Medicine residents	Baseline: Comfort searching MEDLINE and filtered EBM resources. After intervention: OSSE: successfull searches, search time, search technique	Assignment
GP trainees	Appliance of EBM concepts, and effect of answered questions (by interns)	Questionnaire
Internal Medicine residents	Information exchange behaviour in precepting sessions	Observational
Radiation oncology	Extent to internet access and level of MEDLINE use and access to online medical journals, awareness of Cochrane library and other EBM resources	Questionnaire
Several, junior and senior medical staff	Frequency and type of CIAP usage	Questionnaire
Several, faculty members, fellows, residents (of teaching tertiary care hospital)	Importance of resources in daily clinical practice	Questionnaire
GP trainees and GPs	Use of research evidence in audit reports, reasons for use	Observational & interview

GP seemed to more often consult their supervisors (range 52%-94%) than trainees in a hospital-based specialty (6-35%).(18,22,34,39) When resources were categorised according to digital or paper, digital resources seemed to be used less, both by trainees GP (3-21% digital and 14-50% paper, N=3 studies), and trainees in a hospital-based specialty (2-42% digital vs. 31-37% paper, N=3 studies), and this did not change over time.(9,18,22,23,27,39) The use of evidence-based resources differed considerably with percentages of 1-73% for primary EBM resources such as PubMed, and 1-48% for secondary EBM resources, such as Cochrane, reviews, and guidelines. (9,18,21,27,34,39) Only two studies reported on the use of guidelines explicitly and showed that trainees GP performed 1-3% of the searches using guidelines.(18,34)

Timing of searches

The moment searches were performed was reported by eight studies (three from primary care).(12,22,23,25,27,29,32,46) Trainees usually searched during working hours, with GPs and trainees GP searching during patient consultations more frequently than their hospital counterparts.(25) Searching during patient consultations was mostly done outside the consultation room (87%), and much less frequently in the presence of the patient (13%).(22) Green et al. reported that 25% of the queries was pursued the same day, the other 75% within 7-10 days.(27) One study quoted that, if trainees GP searched during patient consultations, they "mostly searched for easily retrievable evidence, like required dosage of medication".(12)

Reasons to search or to refrain from searching

Five studies reported reasons to search or to refrain from searching, of which three included trainees in a hospital-based specialty.(11,12,25,43,46) Most frequently reported reasons to refrain from searching, only reported for trainees in a hospital-based specialty, were lack of time (14-60%, based on two studies), forgetting the query (9-29%, based on two studies), willingness to wait for an answer (30%, one study), no effect expected on clinical management (3-36%, based on two studies), and best guess at answering the question immediately (11%, one study).(9,27)

Time spent on searching

Eleven studies reported on duration of searches, of which seven throughout the process of patient management.(9,22,23,25,27,29,34) Three of those reported data on trainees GP.(22,23,25) Trainees GP generally searched less than five minutes, with one study reporting durations of less than one minute for nearly half of the searches (48%). (22) Trainees in a hospital-based specialty searched on average 15 minutes (95%CI 11-18).(27) Only 16% of their searches was less than 5 minutes, whereas a little more than half (56%) took 6-29 minutes.(25) One questionnaire study reported that trainees in a hospital-based specialty searched longer when no answer was retrieved.(34)

Table 3. Frequencies and percentages of clinical queries per patient seen, searches performed, and answers retrieved.

Study	Patients	Queries (probability)	Searches (probability)	Answers (probability)	Comments
Trainees GP					
Observations					
Allan, 2012	420	325 (0.77)	406 (1.25)	334 (0.83)^	^ Classified as 'no answer' when categorized as 'no help' by trainees
McCord, 2007	328	NA	NA	532	
Ramos, 2003	139	213 (1.5)	165 (0.77)	NR*	
Mean (95%CI)	***296***	***(1.15, 0.41-1.89)***	***(1.01, 0.54-1.48)***	***0.83 (NR)***	
Trainees in a hospital-based specialty					
Observations					
Green, 2000	403	280 (0.69)	80 (0.29)^	NR	^ Interview 7-10 days after each clinic session about 277 (99%) queries
Interviews					
d'Allessandro,1999	NR	182	138 (0.76)	114 (0.83)	
Information service					
Crowley, 2003	NR	625	581** (0.93)	477** (0.82)	Data were obtained during an introduction of a web-based system collecting clinical queries
Coulthard, 2001	NR	NA	41	(31)^ (0.76)	^Data obtained from questionnaire
Search log					
Chisholm, 2012	25665**	NA	18871	NR	
Assignment					
Schilling, 2005	NR	158	NA	141*	
Questionnaire					
Martinez-Silveira, 2005	NR	67	65 (0.97)	49 (0.77)^	^Answers were retrieved partially for 14 (0.22) searches
Mean (95%CI)	***13034***	***(0.69, 0.61-0.78)***	***(0.74, 0.43-1.05)***	***(0.79, 0.75-0.83)***	
Trainees					
Mean (95%CI)	***403 (IQR 234-13043)***	***(1.00, 0.48-1.52)***	***(0.83, 0.57-1.09)***	***(0.79, 0.76-0.82)***	

Abbreviations: NA, not applicable (the study collected and analyzed only the queries that trainees pursued); NR, not reported.
*Not reported for trainees specifically
** Estimated from data reported

Impact of retrieved answers

Fifteen studies reported on the impact of retrieved answers, of which five included trainees GP.(9,11,18,21–23,27,32,34,35,40,44–46,49) Most studies mentioned that retrieved answers did impact on patient management, varying from 22-69% for trainees GP (N=2 studies) (18,21), to 19-88% for trainees in a hospital-based specialty (9,27,35,49). None of the studies measured the impact on patient care in daily clinical practice, but only measured the impact as perceived by the participants.

DISCUSSION

Our systematic review shows a large variety in design, setting, participants, sample size, instruments, and outcome data in the studies reporting on trainees' clinical questions and seeking behaviour in daily practice. Both trainees GP and trainees in a hospital-based specialty frequently encounter clinical queries in daily clinical practice, perform searches in the vast majority, and retrieve answers in most cases. Most queries regarded therapeutic questions. Colleagues were used as their primary source of information. Searches were mostly performed during patient consultations, more often so by trainees GP than trainees in a hospital-based specialty, but not in the presence of the patient. Most queries focussed on checking the validity of a therapeutic decision. Answers were believed to change patient's management frequently.

To our knowledge this is the first review about trainees' clinical question frequency encountered in daily practice and subsequent information seeking behaviour. Prior reviews assessed the clinical question frequencies and information seeking behaviour of experienced doctors, as opposed to trainees, showing heterogeneous data as well.(1,2) Only seven of our 35 studies were included in these earlier reviews. showing that our review provides complementary data.(9,19,23,27,34,45,49)

The number of clinical queries, searches performed, and answers retrieved varied among studies. From the data available it was not possible to determine the reason(s) for this, as several issues varied each time. It is likely that the instruments used to obtain data played an important role, rather than the differences between trainees GP and trainees in a hospital-based specialty, as they differed greatly and some of these methods may have artificially influenced the findings. Direct observation by a researcher in clinical practice may have influenced the performance of the trainees and stimulated or inhibited the number of queries and seeking behaviour.(50,51) In interviews, queries may have been missed when trainees failed to express them. Direct observations and interviews are time-consuming for researchers, thereby minimizing the sample size, whereas self-reports are time-consuming for participants and could easily lead to bias in the reported number of clinical queries.(2) Participants may, for example, only report clinical queries they assume an answer can be pursued for.(3,52,53)

Compared to the earlier published data from experienced doctors, trainees seemed to have more clinical queries and perform more searches, but retrieve just as many answers

and just as infrequently use evidence-based resources, despite the better availability of online EBM-resources throughout the process of patient management.(1,2) This may be due to the fact that colleagues, the primary information source used to pursue an answer to a clinical query, provide answers much more quickly than other sources of information.(54) Second, EBM training programmes may not adequately address the needs of trainees. After all, there is still not much evidence on how to teach EBM best to improve the practice of EBM in daily clinical practice.(6) Third, the supervisor as a role model is important, since an important barrier for trainees can be the attitude of supervisors towards EBM.(54,55) This is most important for trainees GP, since they merely have just one supervisor.(18) Finally, the availability of evidence-based information throughout the process of patient management may not meet trainees' information needs. The impact of the introduction of clinical decision support systems over the last few years on this is unknown.(1)

Most studies in this review assessed only one aspect of the whole search process (having a clinical query, performing a search, or retrieving an answer). To obtain an overall view on trainees' clinical question frequencies and seeking behaviour future studies should obtain information on different aspects, and should provide information on the number of patients, queries, searches, and answers, preferably. Moreover, most data were not obtained throughout the process of patient management, but retrospectively with self-reported surveys. Retrieving information from Electronic Health Records (EHRs) would enable to collect a larger set of objective data from routine clinical practice, and assess the effects of seeking behaviour on (the quality of) patient care. Furthermore, factors contributing to differences in clinical question frequencies and seeking behaviour among trainees have not been assessed yet. Differences in health care systems, education methods, and availability of (electronic) resources between countries may limit the generalizability of our findings. Finally, future studies should assess the effects of EBM training programmes on trainees' clinical question frequencies and seeking behaviour, since this would be instrumental to optimize EBM training programmes.

This systematic review on trainees GP and trainees in a hospital-based specialty' clinical question frequencies encountered in daily practice and seeking behaviour shows that, although trainees pursue answers to the vast majority of clinical queries, the use of (online) evidence-based resources needs improvement. Optimisation of EBM training programmes could establish this. Future research should focus on the actual learning needs of trainees in EBM training programmes, as identification of these needs can be used as a basis to optimize the learning process of trainees.(48,56,57) Optimising the availability of guidelines and aggregated evidence (e.g. Cochrane) throughout the process of patient management and further integration of clinical decision support systems in the EHRs may help to match trainees' information needs better.(1) These support systems could also be applied to facilitate research on seeking behaviour throughout the process of patient management, by extracting data from these systems. At last, the role of supervisors should be considered more explicitly in future research, as they are important role models for trainees.(48)

REFERENCES

1 Del Fiol G, Workman TE, Gorman PN. Clinical questions raised by clinicians at the point of care: a systematic review. JAMA Intern Med. 2014;174(5):710–8.

2 Davies K, Harrison J. The information-seeking behaviour of doctors: a review of the evidence. Health Info Libr J. 2007 Jun;24(2):78–94.

3 Coumou HCH, Meijman FJ. How do primary care physicians seek answers to clinical questions? A literature review. J Med Libr Assoc. 2006 Jan;94(1):55–60.

4 Frank JR. The CanMEDS 2005 physician competency framework. Better standards. Better physicians. Better care. Ottawa R Coll Physicians Surg Canada. 2005;

5 Dawes M, Sampson U. Knowledge management in clinical practice: a systematic review of information seeking behavior in physicians. Int J Med Inform. 2003 Aug;71(1):9–15.

6 Ilic D, Maloney S. Methods of teaching medical trainees evidence-based medicine: A systematic review. Med Educ. 2014;48(2):124–35.

7 Thomson JS, Anderson K, Haesler E, Barnard A GN. The learner's perspective in GP teaching practices with multi-level learners: a qualitative study. BMC Med Educ. 2014;Mar 19(14):55.

8 Ely JW, Osheroff JA, Chambliss ML, Ebell MH, Rosenbaum ME. Answering physicians' clinical questions: obstacles and potential solutions. J Am Med Inform Assoc. 2005 Jan;12(2):217–24.

9 D'Alessandro MP, Nguyen BC, D'Alessandro DM. Information needs and information-seeking behaviors of on-call radiology residents. Acad Radiol. 1999 Jan;6(1):16–21.

10 Cabell CH, Schardt C, Sanders L, Corey GR, Keitz SA. Resident utilization of information technology. J Gen Intern Med. 2001 Dec;16(12):838–44.

11 Martinez-Silveira MS, Oddone N. Information-seeking behavior of medical residents in clinical practice in Bahia, Brazil. J Med Libr Assoc. 2008 Oct;96(4):381–4.

12 Zwolsman SE, van Dijk N, de Waard MW. Observations of evidence-based medicine in general practice. Perspect Med Educ. 2013;2(4):196–208.

13 Chan AKL. Evidence-based medicine for general practitioners - A tool for lifelong self-directed learning? Hong Kong Pract. 2004;26(9):400–7.

14 Christodoulou I. The evaluation of internet as a tool for evidence based medical research. Technol Heal Care. IOS Press; 2005 Dec 1;13(5):369–70.

15 Santoro E, Della Vedova R, Pandolfini C GL. Internet and continuing medical education in general practice. Ric e Prat. 2002;18(2):56–67.

16 Stevermer JJ, Chambliss ML, Hoekzema GS. Distilling the literature: a randomized, controlled trial testing an intervention to improve selection of medical articles for reading. Acad Med. 1999 Jan;74(1):70–2.

17 Marshall J, Mead P, Jones K, Kaba E TP. Implementing evidence-based practice in primary care: perceptions of a multifaceted programme to encourage guideline use. J Clin Excell. 2002;4(2):223–9.

18 Allan GM, Ma V, Aaron S, Vandermeer B, Manca D, Korownyk C. Residents' clinical questions: how are they answered and are the answers helpful? Can Fam physician Médecin Fam Can. 2012 Jun;58(6):e344–51.

19 Bergus GR, Emerson M. Family medicine residents do not ask better-formulated clinical questions as they advance in their training. Fam Med. 2005;37(7):486–90.

20 Grad R, Macaulay AC, Warner M. Teaching evidence-based medical care: description and evaluation. Fam Med. 2001 Sep;33(8):602–6.

21 Korownyk C, Ross S, Ma V, Aaron S, Allan GM. A descriptive analysis of a novel intervention to help residents become evidence users. Med Teach. 2013 Oct;35(10):e1546–50.

22 McCord G, Smucker WD, Selius BA, Hannan S, Davidson E, Schrop SL, et al. Answering questions at the point of care: do residents practice EBM or manage information sources? Acad Med. 2007 Mar;82(3):298–303.

23 Ramos K, Linscheid R, Schafer S. Real-time information-seeking behavior of residency physicians. Fam Med. 2003 Apr;35(4):257–60.

24 Slawson DC, Shaughnessy AF. Teaching information mastery: creating informed consumers of medical information. J Am Board Fam Pract. 1999;12(6):444–9.

25 Brennan N, Edwards S, Kelly N, Miller A, Harrower L, Mattick K. Qualified doctor and medical students' use of resources for accessing information: what is used and why? Health Info Libr J. 2014 Sep;31(3):204–14.

26 Forrest M, Robb M. The information needs of doctors-in-training: case study from the Cairns Library, University of Oxford. Health Libr Rev. 2000 Sep;17(3):129–35.

27 Green ML, Ciampi MA, Ellis PJ. Residents' medical information needs in clinic: are they being met? Am J Med. 2000 Aug 15;109(3):218–23.

28 Tilburt JC, Goold SD, Siddiqui N, Mangrulkar RS. How do doctors use information in real-time? A qualitative study of internal medicine resident precepting. J Eval Clin Pr. 2007;13(5):772–80.

29 Hayward RS, El-Hajj M, Voth TK, Deis K. Patterns of use of decision support tools by clinicians. AMIA Annu Symp Proc. 2006 Jan;329–33.

30 Chisholm R, Finnell JT. Emergency department physician internet use during clinical encounters. AMIA Annu Symp Proc. 2012;2012(March 2010):1176–83.

31 Kim S, Willett LR, Murphy DJ, O'Rourke K, Sharma R, Shea J a. Impact of an evidence-based medicine curriculum on resident use of electronic resources: a randomized controlled study. J Gen Intern Med. 2008 Nov;23(11):1804–8.

32 León S a, Fontelo P, Green L, Ackerman M, Liu F. Evidence-based medicine among internal medicine residents in a community hospital program using smart phones. BMC Med Inform Decis Mak. 2007 Jan;7:5.

33 Stark R, Helenius IM, Schimming LM, Takahara N, Kronish I, Korenstein D. Real-time EBM: From Bed Board to Keyboard and Back. 2007;(April):1656–60.

34 Schilling LM, Steiner JF, Lundahl K, Anderson RJ. Residents' patient-specific clinical questions: opportunities for evidence-based learning. Acad Med. 2005 Jan;80(1):51–6.

35 Coulthard MG, Callaghan LA. An evidence-based medicine fellowship in a children's teaching hospital. J Qual Clin Pract. 2001 Dec;21(4):126–30.

36 Duran-Nelson A, Gladding S, Beattie J, Nixon LJ. Should we Google it? Resource use by internal medicine residents for point-of-care clinical decision making. Acad Med. 2013 Jun;88(6):788–94.

37 Feldstein D a, Maenner MJ, Srisurichan R, Roach M a, Vogelman BS. Evidence-based medicine training during residency: a randomized controlled trial of efficacy. BMC Med Educ. 2010;10:59.

38 Hadley JA, Wall D, Khan KS. Learning needs analysis to guide teaching evidence-based medicine: knowledge and beliefs amongst trainees from various specialities. BMC Med Educ. 2007 Jan;7:11.

39 Jeve YB, Doshani A, Singhal T, Konje J. Knowledge, skills and attitude of evidence-based medicine among obstetrics and gynaecology trainees: a questionnaire survey. JRSM Short Rep. 2013 Dec;4(12):2042533313498719.

40 Phua J, Lim TK. How residents and interns utilise and perceive the personal digital assistant and UpToDate. BMC Med Educ. 2008 Jan;8:39.

41 Risahmawati RR, Emura SS, Nishi TT, Koizumi SS. Japanese Resident Physicians' Attitudes, knowledge, and Perceived Barriers on the Practice of Evidence Based Medicine: a Survey. BMC Res Notes. 2011 Jan;4:374.

42 Sadeghi-Ghyassi F, Nosraty L, Ghojazadeh M, Mostafaie A. Evidence databases application: comparison of university faculties versus clinical residents in a developing country. J Eval Clin Pract. 2013 Apr;19(2):292–7.

43 Shirkhedkar P, Day AS. Medical literature search practice in paediatric junior medical staff: a questionnaire survey. Inform Health Soc Care. 2008 Mar;33(1):69–75.

44 Thom DH, Haugen J, Sommers PS, Lovett P. Description and evaluation of an EBM curriculum using a block rotation. BMC Med Educ. 2004 Oct 11;4:19.

45 Veness M, Rikard-bell G, Ward J. Views of Australian and New Zealand radiation oncologists and registrars about evidence-based medicine and their access to Internet based sources of evidence. 2003;(October 2001):409–15.

46 Westbrook JI, Gosling AS, Westbrook MT. Use of point-of-care online clinical evidence by junior and senior doctors in New South Wales public hospitals. 2005;399–404.

47 Yousefi-Nooraie R, Shakiba B, Mortaz-Hedjri S SH. Sources of knowledge in clinical

practice in postgraduate medical students and faculty members: a conceptual map. J Eval Clin Pr. 2007;13(4):564–8.

48 Zwolsman SE, van Dijk N, Te Pas E, Wieringa-de Waard M. Barriers to the use of evidence-based medicine: knowledge and skills, attitude, and external factors. Perspect Med Educ. 2013;2(1):4–13.

49 Crowley SD, Owens T a, Schardt CM, Wardell SI, Peterson J, Garrison S, et al. A Web-based compendium of clinical questions and medical evidence to educate internal medicine residents. Acad Med. 2003;78(3):270–4.

50 Barrie a R, Ward a M. Questioning behaviour in general practice: a pragmatic study. BMJ. 1997 Dec 6;315(7121):1512–5.

51 Gerrity MS, DeVellis RF, Earp JA. Physicians' reactions to uncertainty in patient care. A new measure and new insights. Med Care. 1990 Aug;28(8):724–36.

52 Arroll B, Pandit S, Kerins D, Tracey J, Kerse N. Use of information sources among New Zealand family physicians with high access to computers. (JFP Online). J Fam Pract. Quadrant Healthcom, Inc.; 2002 Aug 1;51(8):706–7.

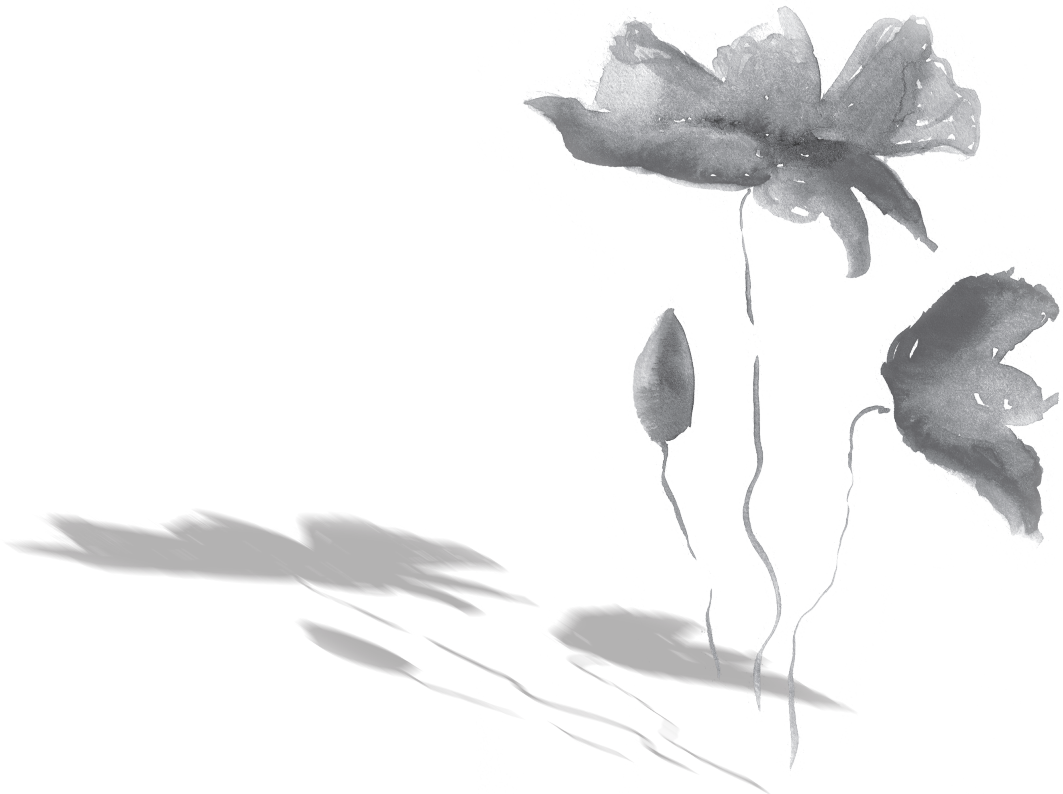
53 Ebell MH, White L. What is the best way to gather clinical questions from physicians? J Med Libr Assoc. 2003 Jul;91(3):364–6.

54 te Pas E, van Dijk N, Bartelink MEL, Wieringa-De Waard M. Factors influencing the EBM behaviour of GP trainers: A mixed method study. Medical Teacher. 2012. p. e1–8.

55 Paice E. How important are role models in making good doctors? BMJ. 2002 Sep 28;325(7366):707–10.

56 Zwolsman SE, Dijk N Van, Verhoeven AAH, Ruijter W De, Waard MW. How learning style affects evidence-based medicine: a survey study. BMC Med Educ. BioMed Central Ltd; 2011;11(1):81.

57 van Dijk N, Hooft L, Wieringa-de Waard M. What are the barriers to residents' practicing evidence-based medicine? A systematic review. Acad Med. 2010;85(7):1163–70.



CHAPTER 5

General practice trainees' information searching strategies for clinical queries encountered in daily practice

Based on: Kortekaas MF, Bartelink ML, Boelman L, Hoes AW, de Wit NJ. General practice trainees' information searching strategies for clinical queries encountered in daily practice. Fam Pract.2015 Oct;32(5):533-7.

Kortekaas MF, Bartelink MEL,
Boelman L, Hoes AW, de Wit NJ

"Het belangrijkst is niet ophouden met vragen stellen."
Albert Einstein

ABSTRACT

Background

Earlier studies have shown that clinical queries are common among doctors. Data on the information seeking behaviour of GP trainees are scarce though, and numbers studied small.

Objective

The objective of this study was to determine how often and how GP trainees search for answers to clinical queries encountered in daily clinical practice.

Methods

Third-year GP trainees kept logs on all patient contacts for eight consecutive practice days. Information was obtained on: patient contacts (description), clinical queries (frequency, type), seeking behaviour (frequency, moment, reason not to search, resources used, duration of search), and answers (frequency, impact). Descriptive analyses were performed; frequencies and percentages were computed. We calculated the number of clinical queries per patient, the number of searches per query, and the number of answers per search.

Results

Seventy-six trainees reported 1,533 clinical queries about 7,300 patients presenting 7,619 complaints (mean of 0.2 queries per patient, SD 0.1). For most of the queries trainees pursued an answer (mean of 0.8 per query, SD 0.2), mostly during consultation (61% of searches), and frequently retrieved answers (mean of 0.8 per search, SD 0.17) they reported to improve clinical decision making in 26%. Most common resources were colleagues or supervisors (28%), and national GP guidelines (26%). The median duration of a search was 4 minutes (IQR 3).

Conclusion

GP trainees have one clinical query per five patients. They often attempt to find answers and reported to succeed in most of the searches, primarily by consulting supervisors or colleagues and national GP guidelines.

INTRODUCTION

General practice (GP) trainees in the Netherlands learn to work according to the principles of evidence-based medicine (EBM). EBM aims at integrating clinical expertise, patient values and the best available clinical evidence in daily clinical practice.(1) EBM is taught according to the five steps as stated in the Sicily Statement; ask, acquire, appraise, apply and assess.(2) An essential skill that trainees learn during their EBM training is to search for evidence-based answers to clinical queries they encounter in daily clinical practice. Trainees are reported to have many clinical queries; 0.7 to 1.6 clinical queries per patient.(3-6)

There are different ways to deal with queries, and different ways to search for answers. The evidence required to answer queries can be ranked based on the quality of the available evidence in the so-called EBM pyramid.(7) According to the described hierarchy of information resources for evidence, GP trainees can use 1) guidelines/pre-appraised resources, 2) unfiltered information sources, and/or 3) expert opinion or background information. To optimize EBM training, information on what type of clinical queries GP trainees have, and how often and how they try to pursue answers would be instrumental.

The objective of the current study is to determine the seeking behaviour of GP trainees, by assessing how often and how they search for answers to clinical queries they encounter in daily clinical practice.

METHODS

Design

We used baseline data from the PINET study (Personalized INtegrated Evidence-based medicine teaching for Trainees in general practice), a prospective, cluster randomized controlled trial, in which third-year GP trainees were allocated to either an integrated EBM training program or a standalone (i.e. the regular) training program for trainees and their supervisors. The primary outcome of the PINET study was EBM behaviour of GP trainees in terms of 1) seeking behaviour (8), and 2) adherence to the practice guidelines of the Dutch College of General Practice.(9) Secondary outcomes were knowledge of and attitude towards EBM.(12)

Setting

Our study was performed within the GP vocational training program of the University Medical Centre Utrecht, the Netherlands, in 2011.

Study population

All third-year GP trainees that participated in the GP vocational training program (*n=82*) were included when they entered their final year of training. Demographics and baseline characteristics of all participants were collected.

Data collection

Trainees were asked to keep logs on all patient consultations during eight consecutive practice days. Telephone contacts with patients were not reported. Data were collected using a paper or digital log, depending on the trainee's preference. In table 1 is shown what data were collected on 1) patient contact information, 2) clinical queries, 3) seeking behaviour, and 4) answers. A clinical query was defined as every patient-related question that was a reason to search for an answer. Seeking behaviour was defined as searching for an answer to a clinical query using all possible human or written information resources. Human resources pertained GP trainers, colleagues or specialists. Written resources pertained textbooks, guidelines (national GP and other guidelines), PubMed, and pre-appraised bibliographic databases (such as Clinical Evidence, TRIP, Cochrane). All written resources but textbooks were considered evidence-based, as textbooks could be out-dated. Trainees were allowed to report more than one complaint per patient contact, and more queries in each consultation. In paper logs trainees could report if they used more than one resource and this was categorized as 'multiple resources'. Trainees reported what impact they thought the retrieved information had on their clinical management decision. As shown in table 1 all questions about 'seeking behaviour', and 'answers to clinical queries' were multiple-choice questions, containing an option 'other'. Trainees were allowed to report more

Table 1. Seeking behaviour in third-year GP trainees (N=76) from the GP vocational training program in Utrecht, in 2011 (data collected from logs).

Item	Variables
Patient contact information	Sex Age Anonymous description of the patient contact, as recorded in the Electronic Medical Recording system (anamnesis, physical examination, consult evaluation, and plan) Reason for encounter (ICPC code)(10)
Clinical query	Presence of clinical query (yes/no) Written description Type of question (diagnosis, aetiology, prognosis, therapy)
Seeking behaviour	Presence of seeking behaviour (yes/no) Duration of search (in minutes) If yes: Search moment (before patient consultation, during patient consultation, directly after patient consultation, later on the same day, at home, other) If no: Reason (little relevance of clinical question, lack of time, pragmatic approach*, other)
Answer to the clinical query	Answer retrieved (yes/no) Resource (attending GP/colleague, national GP guideline, other guideline(s), textbook, consultation of a specialist, pharmacotherapeutic guideline, PubMed, pre-appraised bibliographic databases, other) Impact, according to trainee (improvement of clinical decision making**, confirmation of the decision, expanding knowledge, recall of knowledge, reassurance, none, other)

* Pragmatic approach: based their clinical management on the at that moment available information.
** According to the trainee.

than one complaint per patient contact, and more queries in each consultation. At baseline, all trainees were asked to self-assess their EBM knowledge and EBM attitude on a 5-point Likert scale ranging from very poor to very good.

Data analysis

Data were recorded in SPSS version 20.0. Presented complaints were classified according to the International Classification of Health Problems of Primary Care (ICPC) by the researchers.(10) Frequencies and percentages of the categorical variables were computed. For all normally distributed continuous variables means and standard deviations (SD) were calculated. For non-normally distributed continuous data medians and interquartile ranges (IQR) were calculated. Seeking behaviour characteristics were calculated as 1) number of clinical queries as a proportion of all patient contacts, 2) number of performed searches as a proportion of all queries, and 3) number of retrieved answers as a proportion of all performed searches.

Ethics

This study design was assessed by the UMC Utrecht Ethics Committee, and regarded as non-eligible for full informed consent. However, we obtained informed consent from the participants for the use of the data from questionnaires, tests and logs. All trainees reported patient related data in such a way that these could not be related to individual patients (thus no name, address, etcetera were reported).

RESULTS

Study population

Eighty-two trainees participated. Six trainees were excluded because of changes in their personal training program (*n=1*), maternal leave (*n=2*), and personal reasons (*n=3*). Thus, data of 76 trainees were included in the analysis. In line with the gender distribution in the GP training program trainees were predominantly female (72%), the median age was 31 years (IQR 5). Most (95%) trainees did a hospital internship before starting the GP vocational training program for a median time period of 20 months (IQR 19). Two (3%) trainees had a PhD. Two third (68%) of the trainees worked fulltime during the GP training. Most trainees graded their knowledge of EBM as bad (score 2 of 5, 38%) or neutral (score 3 of 5, 46%). More than half (59%) thought it was important to work according to the principles of EBM.

Patients

Seventy-six trainees reported data on 7,300 patient contacts with 7,619 different reasons for encounter as collected on four to ten practice days per trainee (mean of 7.4 practice days, SD 1.2). Trainees encountered on average a total of 96 patients (SD 23, a mean of 13 patients per day (SD 3)). Most frequent reasons for encounter of the patients were

skin (ICPC category S, 19%), musculoskeletal (L, 18%), or respiratory complaints (R, 17%). Least frequent reasons for encounter of the patients included 'hematologic and lymphatic system' (B, <1%), the male genital system (Y, <1%), and social problems (Z, <1%).

Clinical queries

Trainees reported a total of 1,533 clinical queries for 7,300 patient contacts with a mean of 0.22 clinical queries per patient (SD 0.11, min 0.02, max 0.53). Clinical queries most frequently pertained to the following clinical topics; skin problems (S, 23%), musculoskeletal complaints (L, 15%), and respiratory complaints (R, 13%). Clinical topics that were least frequently encountered by trainees most frequently raised queries; the male genital system (Y, <1% encountered, 39% clinical queries), 'hematologic and lymphatic system' (B, <1% encountered, 35% clinical queries), and 'pregnancy, childbirth, and family planning' (W, 2% encountered, 27% clinical queries). Most frequently clinical queries were therapeutic (50%) or diagnostic (36%) queries; prognostic (5%) or etiologic (9%) queries were less frequent.

Seeking behaviour

For 1,207 of the 1,533 clinical queries trainees attempted to find an answer with a mean of 0.80 searches per clinical query (SD 0.21, min 0.13, max 1.0). Twenty (26%) trainees reported they always searched for an answer when they had a clinical query. Most frequent reported reasons for trainees not trying to find an answer to their query was that they chose a pragmatic approach, i.e. based their clinical management on the available information at that moment (48%), followed by lack of time (18%), the query being considered of little clinical relevance (8%), or other (22%), such as referral, performing additional diagnostic tests, or postponed seeking behaviour. Data on the resources used by trainees to find the answer are shown in table 2. Most frequently used resources were colleagues (35%), national professional GP clinical guidelines (26%), and other guidelines (11%). Primary (0.8%) or pre-appraised bibliographic resources (such as Clinical Evidence, TRIP, Cochrane; 0.5%) were hardly used. For diagnostic queries GP trainees most often used their colleagues, tutors or specialists (37%) as the first information resource, while for therapeutic queries they primarily used the national professional GP guidelines (31%). Most of the time (61%) trainees searched for information during the patient consultation. In 17% of the queries they searched later the same day, or directly after the consultation (11%). Median duration of the search was 4 minutes (IQR 3), with nearly one third (30%) being less than 2 minutes. Just one in ten searches (9%) took more than 10 minutes.

Answers to clinical queries

In the vast majority of the 1,207 searches (n=1,003, 14% of total number of patient contacts) GP trainees found an answer to their clinical query with a mean of 0.83 answers per search (SD 0.17, min 0.27, max 1.0). Nineteen trainees (25%) reported

Table 2. Use of resources to retrieve answers to clinical queries encountered in daily clinical practice, as recorded in logs by third-year GP trainees (N=76) from the GP vocational training program, in 2011.

Type		Resource	Frequency (%)
Human		Attending GP/colleague	333 (28%)
		Consultation of a specialist	86 (7%)
Written	Evidence	National GP guideline (9)	309 (26%)
		Pharmacotherapeutic guideline	71 (6%)
		Other guidelines	55 (5%)
		Pubmed	10 (1%)
		Pre-appraised bibliographic databases (such as Clinical Evidence, TRIP, Cochrane)	6 (1%)
	Other	Textbooks	92 (8%)
		Other resources	107 (9%)
Multiple resources*			138 (11%)
Missing			2 (0%)
Total			**1209 (100%)**

* Only possible in paper logs, as trainees could mark more than one resource on paper, while in digital logs they could only mark one resource.

they always found an answer to a clinical query. When colleagues were used as a resource, answers were retrieved for 77% of the searches. When evidence-based resources were used, an answer was retrieved for 83% of the searches (with a success rate of 91% when the national GP guidelines were used). Trainees considered the answers to queries to have resulted in the following: improved clinical decision making (26%), confirmation of the decision (23%), increase in knowledge (18%), and recall of earlier knowledge (14%). In only a few cases (4%) they reported that the information retrieved did not influence clinical management (table 3). In 93% of the therapeutic queries an answer was found, compared to 74% of the diagnostic queries, 89% of the prognostic queries, and 69% of the etiologic queries.

DISCUSSION

To our knowledge this is the first study assessing seeking behaviour of GP trainees using self-reported logs on such a large number of patient encounters in their daily training practice. (3-6) Our results demonstrate that although GP trainees do not have as many clinical queries per patient as reported earlier (one per five patients), they try to pursue an answer in the vast majority of clinical queries and retrieve an answer for the far most part of performed searches. Trainees reported that the retrieved information resulted in improved clinical decision making in the vast majority of clinical queries. It should be emphasized, however, that we did not assess whether the trainees' clinical decisions improved or just changed. They mainly used human resources to search (one

Table 3. Impact on clinical decision-making of answers retrieved from searches following clinical queries in daily clinical practice, as recorded in logs by third-year GP trainees (N=76) from the GP vocational training program in 2011.

Impact*	Frequency (%)
Improvement of the clinical decision making	261 (25.8%)
Confirmation of the decision	230 (22.7%)
Expanding knowledge	183 (18.1%)
Retrieval of knowledge	139 (13.7%)
Reassurance	96 (9.5%)
No influence	44 (4.4%)
Multiple influences	29 (2.9%)
Other	19 (1.9%)
Missing	10 (1.0%)
Total	**1011 (100%)**

* According to the trainee.

third), followed by evidence-based resources, the national professional GP guidelines (one fourth). Other evidence-based resources found on the internet, such as primary or pre-appraised research resources, were used very infrequently.

Comparison with existing literature

The distribution of reasons for encounter as reported by GP trainees matches the distribution in daily practice in the Netherlands, where skin disease, respiratory and musculoskeletal problems represent the top 3 of clinical domains in consultations. (11,12) While the number of performed searches and retrieved answers was high in our study, the number of clinical queries per patient was relatively low compared to earlier reports. (3-6) It could be an illustration of a poor ability of GP trainees in our study to identify clinical queries, but a good ability to translate queries into answerable questions. It may also be a consequence of selective reporting of clinical queries, negatively influenced by the larger logging requirement when having a clinical query. On the other hand, in earlier studies among GP trainees observational methods were used, which may be an explanation for the differences in results. (3-6) According to the trainees an answer usually resulted in improvement of the clinical decision process, confirming results from earlier studies.(13) Reported reasons for trainees not trying to find an answer to a clinical query were in line with previous studies, with lack of time as an important barrier.(14) The fact that in half of the queries trainees reported to refrain from searching and to make clinical decisions based on the best information available, illustrates the pragmatic way GPs in the Netherlands work in daily clinical practice well.

As primary resource for retrieving an answer to a clinical query GP trainees used colleagues, which is in concordance with previous studies. (3-6) An important reason for GP trainees to consult GP trainers and colleagues is probably the fact that colleagues

and trainers provide answers very quickly, in combination with the direct utility of the answer. Yet it is known that professional advice from a colleague is often more experience based than evidence-based.(15) As second most often used resource GP trainees used the Dutch clinical evidence-based GP guidelines. In the Netherlands, the Dutch College of Family Physicians has produced over a 100 evidence-based guidelines since 1989.(9) These guidelines are easily available to all residents, digitally and on paper, and are well appreciated and frequently used by GPs. An advantage of guidelines is that one of the main barriers of GP trainees to practice EBM is bypassed; i.e. time required to search for information and critically appraise it.(14) As the national GP guidelines in the Netherlands are frequently updated, based on for example the availability of clinically relevant new evidence, we qualified these as a good evidence-based resource in the Netherlands. This may be different in other countries. In some training practices a clinical decision support system may have been available that helped to find the related national GP guideline, based on the allocated ICPC code. Besides colleagues and guidelines, GP trainees often used other research resources, mostly Google. We had expected trainees to use other evidence-based primary or aggregated research resources more often, particularly as most clinical queries were within the therapeutic area and answers to such questions are easiest to retrieve in research resources.(16) On the other hand, it may be that searches in Google have resulted in more relevant articles with better access to free full-text publications, as is reported to be easier using Google than using PubMed. (17) It may be that this type of information seeking behaviour of GP trainees is more effective for daily clinical practice. After all, as is shown in our study and previous studies, trainees primarily seek answers to clinical queries during patient consultations.(6) Only very few times trainees try to answer clinical queries later. This indicates that most clinical queries from GP trainees are relatively easy and quick to answer, as is also reflected in the limited time required for a search and the proportion of answers retrieved.(6)

Strengths and limitations

The major strength of this study is its power: we collected data on a large number and a wide variety of patient contacts, collected by a large number of GP trainees. (5,6) Also, as data were collected in daily clinical practice, these reflect 'real-life' seeking behaviour of GP trainees. Although all trainees were from the same university medical centre, we are convinced that the results are generalizable to the Netherlands (and other countries with similar vocational programs) for several reasons. First, the selection for the vocational training programme takes place at the national level, reducing the chance of regional selection of specific trainees. In addition the organisation and structure of the vocational training program in the Netherlands is comparable across different universities, and accessibility to resources are the same across the country (and beyond). Obviously, seeking behaviour by trainees may vary across countries, because of country specific training programs, different levels of EBM incorporated in

the training and differences in accessibility to the internet and availability of guidelines. The time investment needed for this study, i.e. keeping a log, was substantial for the trainees. This may have led to a loss in commitment by GP trainees, and likely to underreporting of clinical queries. Another reason that GP trainees in our study had less clinical queries per patient than reported in earlier studies might be the difference in the methods applied; we used a log, while in other studies direct observations (3-6) or questionnaires (6) were used. In the logs trainees may have selectively reported only those clinical queries they assumed an answer could be pursued for, as a known predictor of seeking behaviour is the expectation that a clear answer can be found. (18) Such selective reporting seems inherent to our approach. An alternative method, direct observation, on the other hand, is more time-consuming and costly. Such a large number of patient consultations as logged in our study, could never have been observed then. Another approach, administration of questionnaires, leads to more time to reflect, which subsequently may partly artificially raise the (reported) number of clinical queries. Besides, with observations and questionnaires no data would be available about clinical queries a trainee did not try to pursue an answer to, or about postponed seeking behaviour.

CONCLUSION

In conclusion, this study demonstrates that when GP trainees have clinical queries in daily practice, they are able to effectively retrieve an answer in the vast majority of queries, most often by obtaining information from colleagues and the national GP guidelines.

REFERENCES

1 Sackett DL, Rosenberg WMC, Gray J a M, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. 1996. BMJ. 1996 Feb;312(71).

2 Dawes M, Summerskill W, Glasziou P, Cartabellotta A, Martin J, Hopayian K, et al. Sicily statement on evidence-based practice. BMC Med Educ. 2005 Jan 5;5(1):1.

3 Davies K, Harrison J. The information-seeking behaviour of doctors: a review of the evidence. Health Info Libr J. 2007 Jun;24(2):78–94.

4 Green ML, Ciampi M a, Ellis PJ. Residents' medical information needs in clinic: are 7they being met? Am J Med. 2000 Aug 15;109(3):218–23.

5 Allan GM, Ma V, Aaron S, Vandermeer B, Manca D, Korownyk C. Residents' clinical questions: how are they answered and are the answers helpful? Can Fam Physician. 2012 Jun;58(6):e344–51.

6 McCord G, Smucker WD, Selius BA, Hannan S, Davidson E, Schrop SL, et al. Answering questions at the point of care: do residents practice EBM or manage information sources? Acad Med. 2007 Mar;82(3):298–303.

7 ebmpyramid.org [Internet]. Available from: http://www.ebmpyramid.org

8 Oude Rengerink K, Zwolsman SE, Ubbink DT, Mol BWJ, van Dijk N, Vermeulen H. Tools to assess evidence-based practice behaviour among healthcare professionals. Evid Based Med. 2013 Aug;18(4):129–38.

9 nhg.org [Internet]. Available from: https://www.nhg.org

10 Lamberts H WM. ICPC. International Classification of Primary Care. Oxford: Oxford University Press; 1987.

11 JCC Braspenning, FG Schellevis, RPTM Grol (redactie). Tweede Nationale Studie naar ziekten en verrichtingen in de huisartspraktijk. Kwaliteit huisartsenzorg belicht. Utrecht/Nijmegen:NIVEL/WOK, 2004.

12 Van Duppen D, Aertgeerts B, Hannes K, Neirinckx J, Seuntjens L, Goossens F, et al. Online on-the-spot searching increases use of evidence during consultations in family practice. Patient Educ Couns. 2007 Sep;68(1):61–5.

13 Korownyk C, Ross S, Ma V, Aaron S, Allan GM. A descriptive analysis of a novel intervention to help residents become evidence users. Med Teach. 2013 Oct;35(10):e1546–50.

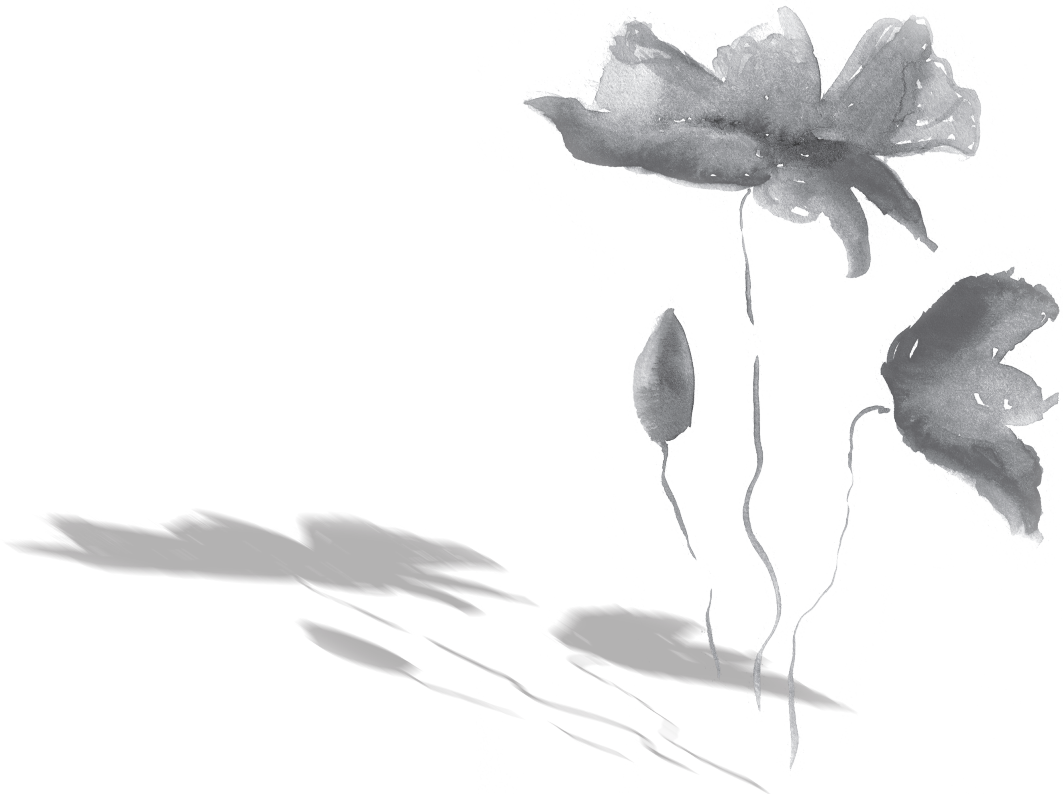
14 Van Dijk N, Hooft L, Wieringa-de Waard M. What are the barriers to residents' practicing evidence-based medicine? A systematic review. Acad Med. 2010 Jul;85(7):1163–70.

15 Schaafsma F, Verbeek J, Hulshof C, van Dijk F. Caution required when relying on a colleague's advice; a comparison between professional advice and evidence from the literature. BMC Health Serv Res. 2005 Jan;5:59.

16 Schilling LM1, Steiner JF, Lundahl K, Anderson RJ. Residents' patient-specific clinical questions: opportunities for evidence-based learning. Acad Med. 2005 Jan;80(1):51-6.

17 Shariff SZ, Bejaimal SA, Sontrop JM, Iansavichus A V, Haynes RB, Weir MA, et al. Retrieving clinical evidence: a comparison of PubMed and Google Scholar for quick clinical searches. J Med Internet Res. 2013 Jan;15(8):e164.

18 Coumou HCH, Meijman FJ. How do primary care physicians seek answers to clinical questions? A literature review. J Med Libr Assoc. 2006 Jan;94(1):55–60.



CHAPTER 6

The Utrecht questionnaire measuring knowledge on clinical epidemiology proved to be valid

Accepted for publication in Journal of Clinical Epidemiology: The Utrecht questionnaire measuring knowledge on clinical epidemiology proved to be valid.

Kortekaas MF, Bartelink MEL, de Groot E,
Korving H, de Wit NJ, Grobbee DE, Hoes AW

"If you are not part of the solution, you are part of the problem."
Eldridge Cleaver

ABSTRACT

Objective

Knowledge on clinical epidemiology is crucial to practice evidence-based medicine (EBM). We describe the development and validation of the Utrecht questionnaire on knowledge on Clinical epidemiology for Evidence-based Practice (U-CEP); an assessment tool to be used in the training of clinicians.

Study design and Setting

The U-CEP was developed in two formats: 2 sets of 25 questions and a combined set of 50.The validation was performed among postgraduate general practice (GP) trainees, hospital trainees, GP supervisors, and experts. Internal consistency, internal reliability (item-total correlation, ITC), item discrimination index (IDI), item difficulty, content validity, construct validity, responsiveness, test-retest reliability, and feasibility were assessed. The questionnaire was externally validated.

Results

Internal consistency was good with a Cronbach's alpha of 0.8. The median ITC and mean IDI were satisfactory. Both sets were perceived as relevant to clinical practice. Construct validity was good. Both sets were responsive, but failed on test-retest reliability. One set took 24, the other 33 minutes to complete, on average. External GP trainees had comparable results.

Conclusion

The U-CEP is a valid questionnaire to assess knowledge on clinical epidemiology, which is a prerequisite for practising EBM in daily clinical practice.

INTRODUCTION

Knowledge of clinical epidemiology is crucial to be able to practice evidence-based medicine (EBM) in daily clinical practice.(1) Practising EBM implies the ability to combine the best available evidence with the clinician's expertise and the patient's preferences.(2) Clinical epidemiology focuses on four important challenges clinicians are faced with. First, how to accurately diagnose a patient's illness (diagnosis, D), second to determine what causes the disease (etiology, E), third how to predict the natural history of the disease in an individual patient (prognosis, P) and fourth to estimate effect of interventions on a patient's prognosis (therapy, Th). In routine clinical practice these four domains are incorporated in medical decision making, following the so called DEPTh model.(1) Clinical epidemiology provides the framework and the knowledge and skills for practitioners to critically appraise research evidence and translate outcomes of research into use in daily clinical practice. Given its importance for adequate evidence-based practising in the future, monitoring theoretical knowledge on clinical epidemiology is important in the training of clinicians.

Testing knowledge needed to practice EBM is essential in clinicians (3), and should focus on those aspects useful in clinical practice. The second Sicily Statement pointed out that for a useful evaluation of EBM training it should be clear which aspect(s) such an assessment instrument intends to measure.(4) A number of questionnaires developed for testing knowledge needed to practice EBM exist already (5–7), but in our view these do not prioritize clinical relevance, are time-consuming to score, or assess therapeutic issues only. Importantly, developers of those questionnaires often provide only minimal data on validation.(5–7)

We previously developed an EBM training program for the vocational training of general practitioners. Focus of the program is the decision process in primary care, and we aim to integrate the training as much as possible into daily clinical practice.(8) The EBM training is strongly based on dilemmas derived from clinical practice and focuses on relevant outcomes for patients. It covers all clinical domains, since many clinical queries pertain not only to therapeutic, but to diagnostic or prognostic topics as well.(8)

We report on the development and validation of the Utrecht questionnaire on knowledge on Clinical epidemiology for Evidence-based Practice (U-CEP), a questionnaire suitable for the evaluation of EBM training, with a focus on those aspects relevant to clinical practice. The U-CEP was developed and validated in Dutch language, but is also available in an English version.

METHODS

Development of the U-CEP

We postulated that an optimal questionnaire should address the content of EBM training, cover as many different aspects of EBM (ask, acquire, appraise, apply, assess) as possible, contain questions on clinically relevant aspects with an equal distribution

across the different types of clinically relevant research (DEPTh), and test the minimal required methodological knowledge to be able to translate research results to clinical practice. At first, we used our experiences as teachers of EBM to include questions on difficulties clinicians frequently encounter in the interpretation of research findings and their use in daily clinical practice. We devised an initial set of 95 items based on the most relevant themes in clinical epidemiology.(1) Experts in the development of summative assessment helped to devise instructions, response options (not including the option "do not know" for example), and rules for scoring.(9) We exchanged opinions about the first drafts between four experienced teachers in EBM (MB, GvdH, MK, and NdW), and adapted elements of the questionnaire accordingly. Finally, the list of 95 items was judged and adapted by two senior clinical epidemiologists (AH, DG). Although there seems to be no agreement on the optimum length of a questionnaire, we aimed to develop a questionnaire that was both as long as needed and as short as possible. We reduced the number of items in the questionnaire on the basis of an item analysis using the scores of respondents. For this we used the data from those who had followed EBM training or were expert on EBM (table 1).

Validation

Population and setting

The validation was performed among 219 postgraduate general practice (GP) trainees (180 first-year and 39 last (i.e. third)-year, 20 hospital trainees, 20 GP supervisors, and 8 expert academic GPs or clinical epidemiologists from the University Medical Center Utrecht (UMCU). Characteristics of participants, such as age, gender, time since graduation, PhD degree, and self-perceived knowledge on EBM were collected through an online survey. The EBM training that the first-year GP trainees and hospital trainees received was a two-day course in which essential skills, such as searching for evidence, critical appraisal of the literature for different research designs, and basic analytic skills were taught in accordance with the five steps of EBM training(10), with focus on clinical relevance and subsequent use by in (future) GPs or specialists in clinical practice.

Procedure

Participants were given access to the questionnaire through an online assessment system, Testvision Online (http://www.testvision.nl). On the first screen the main objectives of the questionnaire and study were explained. While filling in the questionnaire participants were not allowed to use educational material, neither paper versions nor internet based sources. The questions were provided in a random order to reduce any potential topic related disinterest as much as possible and preventing construct-irrelevant order effects that may threaten the validity of the questionnaire. (11) Questionnaires were either filled in just before an EBM training session, just after, or both. We analysed trainees who filled in the questionnaire before and those who filled it in after the training separately. Most participants filled in the questionnaire at

the vocational training centre. When this was not possible (e.g. for the experts not working at the vocational training centre) they filled it in at home.

Validity measures

Individual item analysis

For individual item analysis we analysed the internal consistency (Cronbach's alpha), the item discrimination index (IDI), the internal reliability score (item-total correlation, ITC), and the item difficulty. The internal consistency was defined as the degree to which all questions in the questionnaire measure a single construct.(12) The IDI was defined as the ability of each question to discriminate between those with overall high scores and those with overall low scores.(7,13) It was calculated by separating the participants' total scores into quartiles, and subsequently subtracting the proportion of participants in the bottom quartile who correctly answered each item from the proportion of participants in the top quartile who correctly answered the item.(14) It ranges between -1 and +1. The ITC was defined as the correlation between the question score and the overall score on the questionnaire.(5) The item "difficulty" was defined as the relative difficulty of an item. Per question we calculated the proportion of all participants including the experts (denominator) who answered a question correctly after EBM training (numerator). We considered the questionnaire valid, according to the individual item analysis, when internal consistency was good (Cronbach's alpha >0.7), containing questions with a positive ITC (>0.15) and a positive IDI (>0.2), and no questions answered correctly by less than 10% or more than 90% of the participants after EBM training.(15)

Content validity

Content validity was ascertained by involving experts in EBM training in the development of the questionnaire. We also assessed content validity by asking a convenience sample of 39 third-year GP trainees for their opinion on difficulty and relevance (all on a 5-point Likert scale ranging from 1 ("very easy/poor") to 5 ("very difficult/good")), and the score they expected to have (ranging from 1 ("very poor") to 10 ("very good")). Correlations between expected and achieved scores were calculated using Pearson's correlation. Floor and ceiling effects were assessed as well, defined as the range and distribution of scores.(15) We considered the questionnaire valid on this item when less than 15% of participants achieved the highest or lowest score after EBM training since otherwise no improvement or deterioration of knowledge can be assessed.(15)

Construct validity

Construct validity was defined as the degree to which the scores of the questionnaire were consistent with hypothesized scores.(16) Construct validity was measured by comparing the mean scores of groups of participants whom we expected to score differently on the questionnaire; trainees, supervisors, and experts, both before and after EBM training. Since no differences between GP and hospital trainees were expected, we

analysed them as one group. The three groups were compared using the ANOVA test. Experts were expected to perform best. Trainees were expected to perform significantly better than supervisors, because they are used to the concept of EBM during their undergraduate training programmes in contrast to most GP supervisors. Moreover, we expected GP trainees to have higher perceived 'need to learn'.

Test-retest reliability

Test-retest reliability was defined as the stability of the scores between two moments of filling in the questionnaire. In order to calculate the test-retest reliability, we asked eighteen third year GP trainees participants to fill in the questionnaire twice before EBM training; the second time three months after the first to avoid recall bias. Due to logistical reasons, the first time the questionnaire was filled in on paper. We compared mean scores on the two questionnaires using the paired-samples t test.

Responsiveness

Responsiveness was defined as the ability of the questionnaire to detect change over time in the construct to be measured.(16) We compared the mean scores of participants before EBM training with their mean scores after EBM training. Responses of participants were analysed as one group and separately per group (GP trainees, hospital trainees, and GP supervisors) using the paired-samples t test.

Feasibility

Feasibility of a questionnaire depends on time needed to complete and score a questionnaire. We assessed time to completion (in minutes) and time needed to check and score (in minutes) among a convenience sample of 39 third-year GP trainees.

External validation

The questionnaire was externally validated among GP trainees from another university medical centre, the Academic Medical Centre (AMC) in Amsterdam. Mean scores of GP trainees from the AMC were compared to mean scores of GP trainees from the UMCU using the independent-samples t test.

Statistical analyses were performed using SPSS version 20.0. Results were considered statistically significant at the $p < 0.05$ level. Corrections for multiple testing were introduced to reduce the risk of bias.

RESULTS

Final format of the U-CEP

The shortening process was based on the results of the individual item analysis on internal reliability and consistency, derived from the scores of the respondents (N=154; 49 trainees (29 first-year GP trainees, 20 hospital trainees) and 19 supervisors

after EBM training, 8 experts, and 78 GP trainees who studied medicine at the same university as the postgraduate training program). We started with a 95-item questionnaire and removed the question with the lowest ITC first (-.073). We then recalculated the Cronbach's alpha to check whether the internal consistency remained the same or became better (in our case Cronbach's alpha changed from 0.858 to 0.860). We then removed the next question with the (then) lowest ITC (-.048). This procedure continued until we had to remove a question which would result in an unequal distribution of the questions across the four different clinical domains (DEPTh). Since we aimed at developing two questionnaires and we wanted to have at least two questions per domain per set, the minimum number of questions per domain was four. When the total number of questions per domain would become less than four after removal as based on the (lowest) ITC, we did not remove the question, but continued and removed the next worst question (based on the ITC). This resulted in a removal of 55 questions with ITCs varying between -0.13 and 0.19, resulting in a 40-item questionnaire (Cronbach's alpha 0.90) with three questions with an ITC <0.15. For two domains the minimum amount of four questions was reached. We then looked which questions were removed, and reinserted ten questions we believed were essential to know. The 50 questions were divided over two comparable sets of 25 items each, as this was regarded useful when different sets of questions before and after EBM training were necessary. Results on psychometric properties of the two separate, but comparable, sets of 25 questions each and the combined set (50 questions) are shown in table 1. Below we report on the two sets of 25 questions only. Both sets of the U-CEP were developed and validated in Dutch. These two versions were subsequently translated into English using forward-backward translation and the English versions were checked for inconsistencies by two senior clinical epidemiologists. The English versions have not yet been applied in a teaching setting.

As shown in the supplementary, both sets of 25 questions have 6 open ended and 19 (single-best only) multiple choice questions. Most questions are scored one for a correct answer and zero for an incorrect answer, four open-ended questions are scored zero to three with one for each part of the question; domain (1 point), determinants (1 point), and outcome (1 point). The maximum score for set A is 33 and for set B 34. Both sets cover the different aspects of EBM well, with set A containing 7 questions on 'ask', 14 on 'appraise', and 4 on 'apply'. Set B contains 6 questions on 'ask', 13 on 'appraise', and 6 on 'apply'. The questions are equally distributed across the different clinical domains in both set A (2 diagnostic, 3 etiologic, 3 prognostic, 4 therapeutic questions), and set B (4 diagnostic, 1 etiologic, 2 prognostic, and 4 therapeutic questions). Both sets contain calculation questions and general questions about clinical epidemiology (e.g."What is the most important reason to ask patients for informed consent?").

Table 1. Results table, validity of the U-CEP.

Test property	Measure used	Participants	Acceptable results
Content validity (test covers entire topic of interest)	Expert opinion and survey	4 experts and 39 third-year GP trainees	Questionnaire covers all main aspects of EBM that are relevant to clinical practice.(5)
Construct validity (evidence that the test measures the construct it intends to)	Mean scores (95% CI) of experts, trainees and supervisors, the latter two after EBM training, compared by ANOVA	76 participants: 8 experts, 49 trainees (29 first-year GP trainees, 20 hospital trainees), and 19 supervisors.	Significant difference in mean scores, with experts performing better than trainees, and trainees performing better than supervisors.
Internal consistency (degree to which all test questions on the test measure a single construct)	Cronbach's alpha—average of all possible split half correlations	154 participants: 49 trainees (29 first-year GP trainees, 20 hospital trainees), and 19 supervisors after EBM training, 8 experts, and 78 GP trainees who studied medicine at the same university as the postgraduate training program.	≥0.6-0.7 is considered acceptable, ≥0.7-0.9 is considered good, ≥0.9 is considered perfect.(7)
Internal reliability (the correlation between the question score and the overall score on the questionnaire)	Itemtotal correlation (ITC) using Pearson's product.	154 participants: 49 trainees (29 first-year GP trainees, 20 hospital trainees), and 19 supervisors after EBM training, 8 experts, and 78 GP trainees who studied medicine at the same university as the postgraduate training program.	≥0.15-0.20 is considered satisfactory, ≥0.20-0.40 is considered good, and ≥0.40 is considered excellent.(7)
Item discrimination (ability of each item to discriminate between those with overall high scores and those with overall low scores)	Item discrimination index (IDI) calculated for each item separately.		>0.20-0.40 is considered satisfactory, ≥0.40 is considered high.(7)
Item difficulty (relative difficulty of each item)	% of participants (who answer) answer a question correctly after EBM training, including experts.	76 participants: 8 experts, 49 trainees (29 first-year GP trainees, 20 hospital trainees), and 19 supervisors.	Wide range (10-90%) of difficulties allows a test to be used with both experts and novices.(5)
Floor and ceiling effects (present when more than 15% of the participants achieve highest or lowest possible scores)	% of participants achieving highest or lowest score after EBM training, including experts.	76 participants: 8 experts, 49 trainees (29 first-year GP trainees, 20 hospital trainees), and 19 supervisors.	≤ 15% achieves highest or lowest possible score.(15)

Performance Set A 25 questions (maximum score 33)	Performance Set B 25 questions (maximum score 34)	Performance Set A and B combined 50 questions (maximum score 66)
Revisions based on experts' suggestions and a survey among third-year GP trainees	Revisions based on experts' suggestions and a survey among third-year GP trainees	Revisions based on experts' suggestions
$M_{experts}$ 30.3 (28.9-31.7) $M_{trainees}$ 25.7 (24.4-27.1) $M_{supervisors}$ 19.3 (16.1-22.1)	$M_{experts}$ 30.0 (28.3-31.8) $M_{trainees}$ 25.4 (24.0-26.7) $M_{supervisors}$ 18.7 (15.9-21.6)	$M_{experts}$ 60.3 (357.2-63.3) $M_{trainees}$ 51.1 (48.5-53.7) $M_{supervisors}$ 38.1 (32.8-43.5)
α = .79	α = .80	α = .89
Median 0.22 (IQR 0.14, range 0.01-0.78) 4 questions <0.15 (16%), 4 questions ≥0.15-0.20 (16%), 13 questions ≥0.20-0.40 (52%), 4 questions ≥0.40 (16%)	Median 0.26 (IQR 0.14, range 0.01-0.74) 4 questions <0.15 (16%), 4 questions ≥0.15-0.20 (16%), 13 questions ≥0.20-0.40 (52%), 4 questions ≥0.40 (16%)	Median 0.24 (IQR 0.14, range -0.06-0.82) 7 questions <0.15 (14%), 9 questions ≥0.15-0.20 (18%), 25 questions ≥0.20-0.40 (50%), 9 questions ≥0.40 (18%)
Median 0.35 (IQR 0.23, range 0.13-0.55) 4 questions <0.20 (16%), 14 questions >0.20-0.40 (56%), 7 questions ≥0.40 (28%).	Median 0.43 (IQR 0.25, range 0.17-0.60) 2 questions <0.20 (8%), 10 questions >0.20-0.40 (40%), 13 questions ≥0.40 (52%).	Median 0.37 (IQR 0.22, range 0.10-0.57) 7 questions <0.20 (14%), 23 questions >0.20-0.40 (46%), 20 questions ≥0.40 (40%).
Ranged from 16-100% with 4 questions>90% and one question 100%. Experts: 38-100% GP trainees: 5-100% Hospital trainees: 7-100% Supervisors: 30-100%	Ranged from 37-96% with 4 questions>90%. Experts: 38-100% GP trainees: 21-89% Hospital trainees: 21-97% Supervisors: 25-100%	Ranged from 8-100% with two questions<10%,10 >90%, and one question 100%. Experts: 38-100% GP trainees: 5-100% Hospital trainees: 7-100% Supervisors: 25-100%
Highest score 32 Lowest score 8 0% achieves highest or lowest possible score	Highest score 32 Lowest score 2 0% achieves highest or lowest possible score	Highest score 64 Lowest score 10 0% achieves highest or lowest possible score

Table 1. *Continued*

Test property	Measure used	Participants	Acceptable results
Test-retest reliability (consistency of a test)	Pearson's correlation of mean scores (±SD) of participants who filled in the questionnaire twice without EBM training in the meantime.	18 first-year GP trainees	No significant differences between the mean scores (95%CI) of both groups.
Responsiveness (questionnaire detects change over time in the construct to be measured)	Mean scores (±SD) of participants who filled in the questionnaire before and after EBM training, compared by paired-samples t-test	65 participants: 49 trainees (29 first-year GP trainees, 20 hospital trainees), 16 supervisors.	Significant better mean scores within groups after EBM training compared to before EBM training.
External validity (the capability of a questionnaire to be used in other settings as well)	Comparison of mean scores (±SD) of GP trainees from the UMCU with GP trainees from the AMC, both after their own EBM training, compared by independent-samples t-test.	101 participants: 47 GP trainees from the UMCU, 54 GP trainees from the AMC.	To be externally valid; no significantly differences in mean scores (95%CI) between both groups.

Population

Hundred-fifty participants filled in the questionnaire before EBM training only (149 first-year GP trainees, 1 GP supervisor), 3 GP supervisors filled in the questionnaire after EBM training only, and 65 participants filled in the questionnaire both before and after EBM training (29 first-year GP trainees, 20 hospital trainees, and 16 GP supervisors. Eight experts and 39 third-year GP trainees filled in the questionnaire once. Characteristics of participants are shown in table 2.

Validity measures

Individual item analysis

Cronbach's alpha of 0.79 for set A, and 0.80 for set B. For set A the median ITC was 0.22 (IQR 0.14, range 0.01 to 0.78) and for set B 0.26 (IQR 0.14, range 0.01 to 0.74). No questions had a negative ITC. In both sets four questions (16%) had an ITC below 0.15. The median IDI for set A was 0.35 (IQR 0.23, range 0.13 to 0.55), and for set B 0.43 (IQR 0.25, range 0.17 to 0.60) with four questions in set A below 0.2, and two in set B. Four questions in each set had answer scores >90%, and one question in set A had a score of 100%. No question had a score below 10%.

Performance Set A 25 questions (maximum score 33)	Performance Set B 25 questions (maximum score 34)	Performance Set A and B combined 50 questions (maximum score 66)
Baseline:18.3 (3.0) After 3 months: 24.7 (4.5) Difference 6.4 (2.9-9.9) Pearson's R -.70, *p*.001	Baseline:17.7 (3.5) After 3 months: 22.9 (4.9) Difference 5.2 (1.9-8.5) Pearson's R -.24, *p*.33	Baseline:36.1 (5.6) After: 47.7 (9.2) Difference 11.6 (5.1-18.1) Pearson's R -.53, *p*.02
Results are shown in table 3	Results are shown in table 3	Results are shown in table 3
UMCU: 24.9 (4.6) AMC:23.3 (6.3) Difference 1.5 (-1.0 – 3.2)	UMCU: 24.1 (4.7) AMC:23.0 (4.2) Difference 1.1 (-0.7 – 2.8)	UMCU: 48.9 (9.1) AMC:46.3 (7.4) Difference 2.6 (-0.6 – 5.9)

Content validity

The response rate to the online survey on content validity was 74% (29 third-year GP trainees from the UMCU). Mean scores on difficulty were 3.5 (SD 0.2) and 3.8 (SD 0.2) for set A and B, respectively. A total of 54% participants considered set A neither easy nor difficult (score 3), and 39% considered it difficult (score 4). For set B 38% considered it neither easy nor difficult (score 3), and 50% difficult (score 4). A total of 77% and 63% considered respectively set A and set B relevant to clinical practice (score>3), with mean scores of 3.9 (SD 0.2) for set A and 2.8 (SD 0.2) for set B. Trainees expected to score on average a 6.2 (SD 0.3) for set A, and a 6.0 (SD 0.2) for set B. These expected scores both correlated positively very strong with achieved scores (Pearson's r 0.80 and 0.70, p<.005). There were no floor and ceiling effects, since none of the participants scored the highest or lowest possible score. Therefore, the questionnaire is valid to assess improvement or deterioration of knowledge.

Construct validity

Both sets showed satisfactory construct validity with both before and after EBM training trainees scoring significantly better than supervisors. Comparisons between

Table 2. Baseline characteristics of responding participants (N=173, 79%).

	First-year GP trainees N=144	Third-year GP trainees N=39	Hospital trainees N=11	GP supervisors N=14	Experts N=4	All N=173
Male	21%	23%	36%	64%	50%	74%
Age	28.5 (3.2)**	31.4 (0.52)	30.6 (3.6)	50.8 (7.9)	-	30.6 (7.4)[#]
PhD	6%	3%[$]	18%	7%		7%[#]
PhD student	6%	3%[$]	9%	0%	-	5%[#]
Time since graduation as MD (in years)	3.9 (3.8)	6.3 (0.54)[$]	4.7 (2.6)	18.6 (7.9)*	-	
Self-perceived knowledge on EBM (Likertscale, score 1-5)	Mean 2.4 (SD 0.8)	Mean 2.7 [$] (SD 0.2)	Mean 2.8 (SD 0.8)	Mean 2.6 (SD 0.5)	Mean 3.8 (SD 0.5)	Mean 2.5 (SD 0.8)

*Time since graduation as a GP. ** Data on one GP trainee missing. [$] Data on ten GP trainees missing. [#] Data on four participants missing. ^ Data on three experts missing.

groups before EBM training showed a significant difference in mean score between trainees (N=200) and supervisors (N=17). The 200 trainees had a mean score (95%CI) of 22.1 (21.4-22.9) on set A, and 21.1 (20.3-21.9) on set B, compared to the mean scores of the 17 supervisors of 16.5 (13.2-19.8) and 15.1 (12.3-18.0), respectively. When experts' mean scores on set A and B were compared to mean scores on both sets of participants after EBM training, experts (30.3 (28.9-31.7) and 30.0 (28.3-31.8)) scored significantly better than trainees (25.7(24.4-27.1) and 25.4 (24.0-26.7)) and supervisors (19.3 (16.1-22.1) and 18.7(15.9-21.6)) (table 1).

Test-retest reliability

Results showed a modest performance on test-retest reliability with significantly different mean scores (±SD) between the two measurements for both set A (18.3±3.0 vs. 24.7±4.5, p.001, Pearson's correlation -.70, p.001), and set B (17.7±3.5 vs. 22.9±4.9, p.004, Pearson's correlation -.24, p.33).

Responsiveness

Both sets proved to be responsive with significant higher mean scores after EBM training than before EBM training for all groups (GP trainees, hospital trainees, and supervisors, see table 3).

Feasibility

The mean time (±SD) needed to complete set A was 24 (SD 11) minutes and 33 minutes for set B (SD 11). For both sets checking and scoring took between 1 and 3 minutes per participant, as only the open ended questions (i.e. calculations and questions in which participants had to identify the research question) had to be checked.

Table 3, responsiveness. Comparisons between scores on the U-CEP before and after EBM training of participants (N=65) who filled in the questionnaire before and after EBM training.

	25-item set A				25-item set B				50-item set, combination of set A and set B			
	Before Score (SD)	After Score (SD)	Difference Score (SD)	95% CI	Before Score (SD)	After Score (SD)	Difference Score (SD)	95% CI	Before Score (SD)	After Score (SD)	Difference Score (SD)	95% CI
GP trainee (N=29)	14.4 (6.9)	24.9 (4.7)	10.5 (7.1)	7.8-13.2	13.9 (7.3)	24.8 (4.6)	10.8 (7.8)	7.8-13.8	28.4 (13.9)	49.7 (9.1)	21.3 (14.7)	15.7-26.9
Hospital trainee (N=20)	24.4 (3.6)	26.9 (4.6)	2.5 (2.6)	1.2-3.7	23.6 (3.9)	26.3 (4.9)	2.6 (2.4)	1.5-3.8	48.0 (7.2)	53.1 (9.3)	5.1 (4.4)	3.0-7.2
Supervisor (N=16)	16.8 (6.6)	20.4 (5.3)	3.7 (5.6)	0.7-6.7	15.3 (5.6)	20.1 (4.5)	4.8 (4.9)	2.2-7.5	32.1 (11.6)	40.6 (9.2)	8.5 (9.8)	3.2-13.8

External validation

Both sets proved to be externally valid with no significant differences in mean scores (±SD) when 54 GP trainees from the AMC were compared to 47 GP trainees from the UMCU (set A; 23.3±6.3 vs. 24.9±4.6 (p.07), set B; 23.0±4.2 vs. 24.1±4.7 (p.22), respectively).

DISCUSSION

The U-CEP measures knowledge on clinical epidemiology, focusing on aspects relevant to daily clinical practice. This is important, because it has become clear that practitioners face challenges in incorporating biomedical knowledge from research into practice.(17,18,19) Other questionnaires measuring this knowledge focus on calculating and interpreting biostatistics (Berlin questionnaire), are more time-consuming to fill in and to score, while they consisted of open-ended questions only (Fresno), or focus on the critical appraisal of therapeutic research only (ACE).(5,6,7) However, clinical epidemiology involves four clinically important challenges clinicians are faced with in daily clinical practice (DEPTh).(1,5,6,7) The importance of knowledge about the differences between these four is emphasized by the presence of different guidelines for complete and accurate reporting for the different types of research.(20,21) Therefore, the U-CEP measures knowledge on all clinical epidemiology aspects. Moreover, the U-CEP is responsive to change. As a result, it is valid to monitor changes in knowledge of clinicians on clinical epidemiology after EBM training. For two other questionnaires, the Fresno and ACE, this has not been assessed.(5,7)

Several aspects of the performance of the U-CEP must be considered critically. First, the third-year GP trainees considered set B of the U-CEP to be of only moderate clinical relevance. As no data on "clinical relevance" are available for earlier EBM

questionnaires, interpretation of these results is difficult.(5,6,7) A possible explanation for this somewhat disappointing score is a different interpretation of the 'relevance for clinical practice' between GP trainees and experienced clinicians and teachers. Possibly, the third-year GP trainees considered the topics dealt with (focusing on diagnosis, prognosis and therapy) in the questionnaire clinically relevant, but the questions themselves (including clinical epidemiology methodology and jargon) were not considered to be clinically relevant. We, however, did not test whether this indeed was the case. Secondly, the U-CEP performs modest on the test-retest reliability with scores at the second measurement comparable to those of GP trainees who received training in EBM. This can be explained by the fact that (for external reasons) the first time distribution was on paper and the second time online, whereas it is recommended that the administration format should remain the same across assessment points. (22) Also, the time-interval of three months may have been too long. In these three months, participating first-year GP trainees may have gained knowledge on clinical epidemiology by self-study or in clinical practice. All GP trainees are supposed to have a textbook on EBM at the start of their GP training programme, and some may have studied this book in preparation of the EBM training, especially in case of disappointing scores at the baseline measurement (explaining the negative correlation in set A). Third, the internal reliability (ITC) and ability to discriminate between participants (IDI) is quite low, albeit this was the case for a small number of questions only.. Removal of these questions did not change the internal consistency of the questionnaire, however, and we consider content validity to be better with those questions included in the questionnaire. Furthermore, although external validation among GP trainees from another university medical centre in the Netherlands proved the U-CEP to be externally valid, further validation of the U-CEP across different undergraduate and postgraduate training programs among other health care professionals and trainees. Validation after translation into other languages will provide valuable information on the validity of the questionnaire when used in other countries, as results may be different.(23) Finally, although it is custom to refer to questionnaires like the U-CEP as questionnaires on 'EBM knowledge (5,6,7), these instruments only assess knowledge, and do not measure whether this knowledge is indeed used in daily clinical practice (and thus leads to evidence-based practice by integrating the three aspects of EBM). Other instruments are needed to assess the latter.

The U-CEP was shown to be valid among different clinicians, such as GP trainees, hospital trainees, and GP supervisors. The online availability allows easy distribution with questions in random order. We have shown that the questionnaire can be used in a valid and reliable way as an evaluation tool for EBM training. Different sets of 25 questions may be used before and after EBM training or one larger set of 50 questions can be used when one wants to use the questionnaire for a formative assessment, to help clinicians learn.

REFERENCES

1 Grobbee D, Hoes A. Clinical Epidemiology: Principles, Methods and Applications for Clinical Research. 1st ed. London: Jones and Bartlett Publishers. 2009.

2 Sackett DL, Rosenberg WMC, Gray J a M, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. 1996. BMJ. 1996 Feb;312(71).

3 Novack L, Jotkowitz A, Knyazer B, Novack V. Evidence-based medicine: assessment of knowledge of basic epidemiological and research methods among medical doctors. Postgrad Med J. England; 2006 Dec;82(974):817–22.

4 Tilson JK, Kaplan SL, Harris JL, Hutchinson A, Ilic D, Niederman R, et al. Sicily statement on classification and development of evidence-based practice learning assessment tools. BMC Med Educ [Internet]. BioMed Central Ltd; 2011;11(1):78.

5 Ramos KD, Schafer S, Tracz SM. Validation of the Fresno test of competence in evidence based medicine. BMJ. England; 2003 Feb;326(7384):319–21.

6 Fritsche L, Greenhalgh T, Falck-Ytter Y, Neumayer H-H, Kunz R. Do short courses in evidence based medicine improve knowledge and skills? Validation of Berlin questionnaire and before and after study of courses in evidence based medicine. BMJ. England; 2002 Dec;325(7376):1338–41.

7 Ilic D, Nordin R Bin, Glasziou P, Tilson JK, Villanueva E. Development and validation of the ACE tool: assessing medical trainees' competency in evidence based medicine. BMC Med Educ [Internet]. 2014 Jan;14:114.

8 Kortekaas MF, Bartelink MEL, Zuithoff NPA, vd Heijden GJMG, De Wit NJ, Hoes AW. Does integrated EBM training in the GP specialty training improve EBM behaviour in daily clinical practice? A cluster randomized controlled trial. Submitted.

9 Ravesloot CJ, Van der Schaaf MF, Muijtjens AMM, Haaring C, Kruitwagen CLJJ, Beek FJA, et al. The don't know option in progress testing. Adv Health Sci Educ Theory Pract. Netherlands; 2015 Dec;20(5):1325–38.

10 Dawes M, Summerskill W, Glasziou P, Cartabellotta A, Martin J, Hopayian K, et al. Sicily statement on evidence-based practice. BMC Med Educ. 2005 Jan 5;5(1):1.

11 Marentette BJ, Meyers LS, Hurtz GM, Kuang DC. Order effects on situational judgment test items: A case of construct-irrelevant difficulty. Int J Sel Assess. 2012;20(3):319–32.

12 Nunnally J. Psychometric theory. Michigan: McGraw-Hill; 1978.

13 Varma S. reliminary item statistics using point-biserial correlation and p-values.

14 Aiken L. Relationships between the item difficulty and discrimination indexes. Educ Psychol Meas. 1979;39:821–4.

15 Zwolsman SE, Wieringa-de Waard M, Hooft L, van Dijk N. Measuring evidence-based medicine knowledge and skills. The Dutch Berlin Questionnaire: translation and validation. Journal of clinical epidemiology. United States; 2011. p. 928–30.

16 Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol [Internet]. 2010 Jul;63(7):737–45.

17 Silva SA, Charon R, Wyer PC. The marriage of evidence and narrative: scientific nurturance within clinical practice. J Eval Clin Pract. England; 2011 Aug;17(4):585–93.

18 Richardson WS. We should overcome the barriers to evidence-based clinical diagnosis! J Clin Epidemiol. England; 2007 Mar;60(3):217–27.

19 Perneger TV1, Ricou B, Boulvain M, Bovier PA, Herrmann FR, Perrier A, Burnand B. Medical researchers evaluate their methodological skills. J Clin Epidemiol. 2004 Dec;57(12):1323-9.

20 Schulz KF, Altman DG, Moher D. CONSORT 2010 Statement: Updated guidelines for reporting parallel group randomised trials. J Clin Epidemiol. United States; 2010 Aug;63(8):834–40.

21 Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ. England; 2015;351:h5527.

22 Thoren ES, Andersson G, Lunner T. The use of research questionnaires with hearing impaired adults: online vs. paper-and-

pencil administration. BMC Ear Nose Throat Disord. England; 2012;12:12.

23 Zwolsman SE, Wieringa-de Waard M, Hooft L, van Dijk N. Measuring evidence-based medicine knowledge and skills. The Dutch Berlin Questionnaire: translation and validation. J Clin Epidemiol. 2011 Aug;64(8):928-30. doi: 10.1016/j.jclinepi.2011.02.005.

APPENDIX 1. U-CEP – SET A - ENGLISH

Questions 1 – 3

Below are five clinical questions. For each question, tick the appropriate box in the table underneath. Choose only one category per question.

1. What is the effect of antibiotics compared to a "watchful waiting" approach on the duration of symptoms in children with acute otitis media?
2. What is the effect of hip surgery compared to pain relief on the quality of life in elderly patients (> 75 years) with coxarthrosis?
3. Is there a causal link between gastroesophageal reflux disease and the occurrence of a chronic cough?

	Diagnostic	Etiologic	Prognostic	Therapeutic
Question 1				
Question 2				
Question 3				

Questions 4 – 5

Which of the research types below study causality?

	Yes	No
4. Etiologic research		
5. Prognostic research		

Question 6

Is it possible to use a panel of experts by means of reference test when the reference test doesn't come close to a 'gold standard'?

☐ Yes
☐ No

Question 7

Is the following statement correct or incorrect?
"Measuring the occlusion rate of the coronary vessels is a relevant outcome in a prognostic study in patients with myocardial infarction of the heart".

☐ Correct
☐ Incorrect

Question 8

Statement: Confounding (bias) plays a role in diagnostic research.

☐ Correct
☐ Incorrect

Question 9
Etiologic research is an example of descriptive research.

- ☐ Correct
- ☐ Incorrect

Question 10
Statement: The objective of randomization is to be able to achieve comparability of baseline patient characteristics.

- ☐ Correct
- ☐ Incorrect

Questions 11-12
In a study, the effects on the occurrence of a stroke are compared in individuals with and without preventive medication (intervention group vs. control group).
In the intervention group (1,000 people) a stroke occurs 50 times within 5 years, in the control group (also 1,000 people) it happens 100 times.

Question 11
What is the relative risk of stroke when comparing the intervention group with the control group?

Question 12
Calculate the number needed to treat for 5 years to prevent one stroke.

Question 13
A confounder is a factor that has a link with the determinant.

- ☐ Correct
- ☐ Incorrect

Question 14
Statement: Blinding of patients and researchers in a trial results in comparability of the external effects.

- ☐ Correct
- ☐ Incorrect

Question 15
What is the relative risk when two treatments are equally effective?

- ☐ 0
- ☐ 1
- ☐ >1
- ☐ >0 and <1

Question 16

What measure of association can usually be calculated on a case-control study?

- ☐ Attributable risk (AR)
- ☐ Number needed to treat (NNT)
- ☐ Odds ratio (OR)
- ☐ Relative risk (RR)

Question 17

Statement: A so-called case-control study is often used to provide insight into the best therapy for a particular disease.

- ☐ Correct
- ☐ Incorrect

Question 18

What is inter-observer variation?

- ☐ The extent to which the observers' interpretations differ.
- ☐ The extent to which observations of a symptom differ from previous observations.
- ☐ A measure for the number of observers that take part in the study.

Question 19

A research bureau is testing a new thermometer. Under the same circumstances it repeatedly gives the same reading, but this is always one degree Celsius too high. Which statement is true about the thermometer?

- ☐ The thermometer is valid and precise.
- ☐ The thermometer is precise and not valid.
- ☐ The thermometer is valid and not precise.
- ☐ The thermometer is neither valid nor precise.

Question 20

A study researches whether it would be beneficial for patients with asthma if their GP was strict in following the treatment guidelines of the Dutch College of General Practitioners. Fifty GPs are randomly divided into two groups. One group of GPs are encouraged in various ways to follow the treatment guidelines as closely as possible. The other group of GPs are left to their own devices (i.e. "usual care"). In all participating practices all known patients with asthma are identified and followed for two years. During that peroid absenteeism from work or school of these patients is documented accurately. What study design is this?

- ☐ Case-control study
- ☐ Prospective follow-up study
- ☐ Randomized controlled trial
- ☐ Retrospective follow-up study

Question 21

A randomized study compares two blood pressure lowering drugs. A calculation has shown that a total of 120 patients are needed in order to detect a difference in the reduction of blood pressure of 5 mmHg with a power of 80 percent. What happens to the required number of patients if you would like to detect a difference of 10 mmHg with the same power?

- ☐ The required number of patients increases.
- ☐ The required number of patients decreases.
- ☐ The required number of patients stays the same

Questions 22-25

Below you will find 4 clinical questions. For each question give the domain, i.e. indicate to which patients / people the results apply. Also give the determinant(s) and outcome. Please note: not every clinical question includes all the components.

Question 22

What is the added value of assessing NT-proBNP to exclude or detect heart failure in patients with shortness of breath on exertion?

Question 23

What is the predictive value of the duration and height of fever in children with a history of bacterial meningitis on their school results?

Question 24

What is the effect of the use of antibiotics compared to "watchful waiting" on the occurrence of an infection after a dog bite?

Question 25

What is the effect of hip surgery compared to pain relief on the quality of life of elderly patients (> 75 years) with coxarthrosis?

APPENDIX 2. U-CEP –SET A - DUTCH

Vraag 1-3

Hieronder staan 3 klinische vragen. Noteer per vraag tot welke categorie deze behoort. Gebruik daarvoor onderstaande tabel en zet een kruis in het juiste vak. Per vraag is slechts 1 categorie mogelijk.

1. Wat is het effect van antibiotica in vergelijking met expectatief beleid op de totale ziekteduur bij kinderen met otitis media acuta?
2. Wat is het effect van een heupoperatie in vergelijking met pijnstilling op de kwaliteit van leven bij oudere patiënten (>75 jaar) met coxartrose?
3. Bestaat er een causaal verband tussen refluxziekte en het optreden van een chronische hoest?

	Diagnostisch	Etiologisch	Prognostisch	Therapeutisch
Vraag 1				
Vraag 2				
Vraag 3				

Vraag 4-5

In welke van onderstaande onderzoeken wordt causaliteit bestudeerd?

	Wel	Niet
4. Etiologisch onderzoek		
5. Prognostisch onderzoek		

Vraag 6

Kan een expertpanel worden gebruikt als referentietest indien de referentietest verre van een "gouden standaard" is?

☐ juist
☐ onjuist

Vraag 7

Stelling: Het meten van het occlusie-percentage van de coronairvaten is een goede uitkomstmaat in een prognostisch onderzoek bij patiënten met een voorwandinfarct van het hart.

☐ juist
☐ onjuist

Vraag 8

Stelling: Confounding (bias) speelt een rol in diagnostisch onderzoek.

☐ juist
☐ onjuist

Vraag 9

Stelling: Etiologisch onderzoek is een voorbeeld van beschrijvend (descriptief) onderzoek.

- ☐ juist
- ☐ onjuist

Vraag 10

Stelling: Randomisatie heeft als doel de baselinekarakteristieken van patiënten vergelijkbaar te maken.

- ☐ juist
- ☐ onjuist

Vraag 11-12

In een onderzoek worden de effecten op het optreden van een CVA met en zonder preventief medicijn (medicijngroep versus placebogroep) vergeleken.In de medicijngroep (1000 personen) komt 50 maal een CVA voor binnen 5 jaar, in de placebogroep (ook 1000 personen) 100 keer.

Vraag 11

Wat is het relatieve risico op CVA wanneer je de interventiegroep vergelijkt met de controlegroep?

Vraag 12

Bereken het number needed to treat (voor 5 jaar).

Vraag 13

Stelling: Een confounder is een factor die verband heeft met de determinant.

- ☐ juist
- ☐ onjuist

Vraag 14

Stelling: Blindering van patiënten en onderzoekers in een trial zorgt voor vergelijkbaarheid van externe effecten.

- ☐ juist
- ☐ onjuist

Vraag 15

Wat is het relatief risico als twee behandelingen even effectief zijn?

- ☐ 0
- ☐ 1
- ☐ >1
- ☐ >0 en <1

Vraag 16

Welke maat kan meestal worden berekend bij een patiënt-controle onderzoek?

- ☐ Attributief risico (AR)
- ☐ Number needed to treat (NNT)
- ☐ Odds ratio (OR)
- ☐ Relatief risico (RR)

Vraag 17

Wordt een zogenaamd patiënt-controle onderzoek vaak gebruikt om inzicht te geven in de beste therapie voor een bepaalde ziekte?

- ☐ juist
- ☐ onjuist

Vraag 18

Wat is interobserver variatie?

- ☐ De mate waarin de interpretatie van waarnemers verschillen vertoont
- ☐ De mate waarin waarnemingen van een verschijnsel verschillen van voorafgaande waarnemingen
- ☐ Een maat voor een aantal verschillende waarnemers dat deelneemt aan het onderzoek

Vraag 19

Een onderzoeksbureau test een nieuwe thermometer. De meter geeft bij herhaalde toepassing in dezelfde omstandigheden dezelfde waarde aan, maar deze is steeds 1 graad. Celsius te hoog. Welke bewering over de thermometer is waar?

- ☐ De thermometer is valide en precies
- ☐ De thermometer is precies en niet valide
- ☐ De thermometer is valide en niet precies
- ☐ De thermometer is niet valide en niet precies

Vraag 20

Onderzocht wordt of patiënten met astma er baat bij hebben als hun huisarts strikt de behandelrichtlijnen van het Nederlands Huisarts Genootschap volgt. Vijftig huisartsen worden willekeurig in twee groepen verdeeld. De ene groep huisartsen wordt op diverse manieren gestimuleerd om zo strikt mogelijke de behandelrichtlijnen te volgen. De andere groep huisartsen wordt daar helemaal vrij in gelaten. In alle deelnemende praktijken worden alle bekende patiënten met astma in kaart gebracht. Gedurende twee jaar wordt het ziekteverzuim van de patiënten nauwkeurig gedocumenteerd. Wat voor onderzoeksopzet is dit?

- ☐ Patiënt-controle-onderzoek
- ☐ Prospectief follow-up onderzoek
- ☐ Randomised controlled trial
- ☐ Retrospectief follow-up onderzoek

Vraag 21

In een gerandomiseerd onderzoek worden twee bloeddrukverlagers met elkaar vergeleken. Men heeft uitgerekend dat er in totaal 120 patiënten nodig zijn om een verschil in bloeddrukdaling van 5 mmHg te detecteren met een onderscheidingsvermogen (power) van 80 procent. Wat verandert er aan het benodigde aantal patiënten als men een verschil van 10 mmHg zou willen kunnen detecteren met hetzelfde onderscheidingsvermogen?

- ☐ Het aantal benodigde patiënten wordt groter
- ☐ Het aantal benodigde patiënten wordt kleiner
- ☐ Het aantal benodigde patiënten blijft gelijk

Vraag 22 tm 25

Hieronder staan 4 klinische vragen. Geef per vraag aan wat het domein is, d.w.z. voor welke patiënten/mensen zijn de resultaten van toepassing. Doe ditzelfde voor determinant en uitkomst. Belangrijk: niet bij elke klinische vraag is elk onderdeel aanwezig.

Vraag 22

Wat is de toegevoegde waarde van inzetten van pro-BNP voor het uitsluiten of aantonen van hartfalen bij patiënten met kortademigheid bij inspanning?

Vraag 23

Wat is de voorspellende waarde van de duur en hoogte van koorts bij kinderen met een doorgemaakte bacteriële meningitis op schoolprestaties?

Vraag 24

Wat is het effect van het gebruik van antibiotica in vergelijking met expectatief beleid op het optreden van een infectie na een hondenbeet?

Vraag 25

Wat is het effect van een heupoperatie in vergelijking met pijnstilling op de kwaliteit van leven bij oudere patiënten (>75 jaar) met coxartrose?

APPENDIX 3. U-CEP –SET B- ENGLISH

Questions 1 – 2

Below are five clinical questions. For each question, tick the appropriate box in the table underneath. Choose only one category per question.

1. What is the effect of the use of antibiotics compared to a "watchful waiting" approach on the occurrence of an infection after a dog bite?
2. What is the added value of assessing NT-proBNP levels to exclude or detect heart failure in patients with shortness of breath on exertion?

	Diagnostic	Etiologic	Prognostic	Therapeutic
Question 1				
Question 2				

Questions 3 – 4

Which of the research types below study causality?

	Yes	No
3. Diagnostic research		
4. Therapeutic research		

Question 5

Is it possible to use a combination of different tests as a reference test when the reference test doesn't come close to a 'gold standard'?

- ☐ Yes
- ☐ No

Question 6

How can prognostic research results best be represented?

- ☐ Relative risk
- ☐ Absolute risk
- ☐ Odds ratio
- ☐ Risk difference

Question 7

Statement: Confounding (bias) plays a role in etiologic research.

- ☐ Correct
- ☐ Incorrect

Question 8

Statement: Prognostic research is an example of descriptive research.

- ☐ Correct
- ☐ Incorrect

Question 9

Statement: Therapeutic research is cross-sectional research.

- ☐ Correct
- ☐ Incorrect

Question 10

In a study, the effects on the occurrence of a stroke are compared in individuals with and without preventive medication (intervention group vs. control group).
In the intervention group (1,000 people) a stroke occurs 50 times within 5 years, in the control group (also 1,000 people) it happens 100 times.

Calculate the risk difference.

Question 11

A diagnostic test is applied to 1,000 patients, 30 of whom have condition X. Sensitivity and specificity of the test for condition X are 80% and 90%, respectively. Which percentage comes closest to the positive predictive value?

- ☐ 20%
- ☐ 40%
- ☐ 60%
- ☐ 80%

Question 12

Statement: A confounder is a factor that has a link with the outcome.

- ☐ Correct
- ☐ Incorrect

Question 13

Statement: In diagnostic examination one is particularly interested in the sensitivity and specificity of the test that is being studied.

- ☐ Correct
- ☐ Incorrect

Question 14

Statement: A study domain is the basis for the generalizability of the results.

- ☐ Correct
- ☐ Incorrect

Question 15

Statement: Blinding of patients and researchers in a trial results in comparability of the baseline differences.

- ☐ Correct
- ☐ Incorrect

Question 16

Statement: A so-called case-control study is often used to provide insight into the etiology of a particular disease.

- ☐ Correct
- ☐ Incorrect

Question 17

In a study of the relationship between alcohol consumption and the incidence of myocardial infarction, a group of patients who had a heart attack in the past year is compared with a control-group of healthy people. The odds ratio that is found for drinking more than two glasses of alcohol per day is 1.42 with a 95% confidence interval from 0.96 to 2.10. What do you conclude about the relationship between the consumption of more than two glasses of alcohol and the incidence of a myocardial infarction?

- ☐ A statistically significant positive relationship. The P value is less than 0.05.
- ☐ A statistically significant positive relationship. The P value is more than 0.05.
- ☐ No statistically significant positive relationship. The P value is less than 0.05.
- ☐ No statistically significant positive relationship. The P value is more than 0.05.

Question 18

A certain study looked at the protective effect of ACE inhibitors on the decline of renal function in patients with diabetes. At a diabetes clinic, a random sample was taken from a group of diabetics in which deterioration of renal function had been established during two years. Prior to that, the prescription medication (ACE inhibitors) was looked at, both for those with relatively *rapid* deterioration of renal function as for those people with relatively *slow* deterioration of renal function during the two year period. What study design is this?

- ☐ Case-control study
- ☐ Prospective follow-up study
- ☐ Randomized controlled trial
- ☐ Retrospective follow-up study

Question 19

A cohort study researches whether there is a relationship between smoking and the development of a particular disease. The results are shown below.

	Disease	
	Present	**Absent**
Smoker	240	760
Non-smoke	260	1740

Calculate the relative risk of developing the disease for a smoker compared to a non-smoker.

Risk category (DVT risk score)	Patients*	DVT present†	DVT absent‡
Very low (0-3)	23	0.7	99.3
Low (4-6)	5	4.5	95.5
Moderate (7-9)	51	21.7	78.3
High (10-13)	21	51.3	48.7

Question 20

Please study the table below.
Prevalence of DVT (deep vein thrombosis) in 4 risk categories when using the diagnostic rule with D-dimer (%)
Statement: The use of the DVT risk score in the table above is particularly useful in the detection of DVT.

- ☐ Correct
- ☐ Incorrect

Question 21

A study on the added value of the erythrocyte sedimentation rate (ESR) in patients with suspected pneumonia in general practice shows that the area under the ROC curve of a diagnostic model with data from history taking, physical examination, and ESR is 0.81. The ROC of a model with the same data from history taking and physical examination but with C-reactive protein (CRP) instead of ESR is 0.90. Which test has the most added value to history taking and physical examination in patients with suspected pneumonia?

- ☐ ESR
- ☐ CRP
- ☐ They are equally valuable.
- ☐ Neither of the tests is adding any value

Questions 22-25

Below you will find 4 clinical questions. For each question give the domain, i.e. indicate to which patients / people the results apply. Also give the determinant(s) and outcome. Please note: not every clinical question includes all the components.

Question 22

What is the effect of antibiotics compared to "watchful waiting" on the total duration of illness in children with acute otitis media?

Question 23

Is it possible that a measles infection causes a higher incidence of inflammatory bowel disease in children?

Question 24

Is there a causal relationship between gastroesophageal reflux disease and the occurrence of a chronic cough?

Question 25

What is the effect of performing an ECG in patients with chest pain in general practice on the frequency or hospital referral?

APPENDIX 4. U-CEP - SET B - DUTCH

Vraag 1-2

Hieronder staan 2 klinische vragen. Noteer per vraag tot welke categorie deze behoort. Gebruik daarvoor onderstaande tabel en zet een kruis in het juiste vak. Per vraag is slechts 1 categorie mogelijk.

1. Wat is het effect van het gebruik van antibiotica in vergelijking met expectatief beleid op het optreden van een infectie na een hondenbeet?
2. Wat is de toegevoegde waarde van inzetten van NTpro-BNP voor het uitsluiten of aantonen van hartfalen bij patiënten met kortademigheid bij inspanning?

	Diagnostisch	Etiologisch	Prognostisch	Therapeutisch
Vraag 1				
Vraag 2				

Vraag 3-4

In welke van onderstaande onderzoeken wordt causaliteit bestudeerd?

	Wel	Niet
3. Diagnostisch onderzoek		
4. Therapeutisch onderzoek		

Vraag 5

Kan een combinatie van verschillende testen worden gebruikt als referentietest indien de referentietest verre van een "gouden standaard" is?

- ☐ juist
- ☐ onjuist

Vraag 6

Hoe kunnen de resultaten bij prognostisch onderzoek het beste worden uitgedrukt?

- ☐ Als relatief risico
- ☐ Als absoluut risico
- ☐ Als odds ratio
- ☐ Als risicoverschil

Vraag 7

Stelling: Confounding (bias) speelt een rol in etiologisch onderzoek.

- ☐ juist
- ☐ onjuist

Vraag 8

Stelling: Prognostisch onderzoek is een voorbeeld van beschrijvend (descriptief) onderzoek.

- ☐ juist
- ☐ onjuist

Vraag 9

Stelling: Therapeutisch onderzoek is dwarsdoorsnede-onderzoek.

- ☐ juist
- ☐ onjuist

Vraag 10

In een onderzoek worden de effecten op het optreden van een CVA met en zonder preventief medicijn (medicijngroep versus placebogroep) vergeleken.
In de medicijngroep (1000 personen) komt 50 maal een CVA voor binnen 5 jaar, in de placebogroep (ook 1000 personen) 100 keer. Bereken het risicoverschil (absolute risicoreductie/attributief risico).

Vraag 11

Een diagnostische test wordt toegepast bij 1000 patiënten van wie er 30 aandoening X hebben. Sensitiviteit en specificiteit van de test voor de aandoening X zijn respectievelijk 80% en 90%. Waar ligt de positief voorspellende waarde het dichtst bij?

- ☐ 20%
- ☐ 40%
- ☐ 60%
- ☐ 80%

Vraag 12

Stelling: Een confounder is een factor die verband heeft met de uitkomst.

- ☐ juist
- ☐ onjuist

Vraag 13

Stelling: In diagnostisch onderzoek is men vooral geïnteresseerd in de sensitiviteit en specificiteit van de test die men bestudeert.

- ☐ juist
- ☐ onjuist

Vraag 14

Stelling: Het domein van een studie vormt de basis voor de generaliseerbaarheid van de resultaten.

- ☐ juist
- ☐ onjuist

Vraag 15

Stelling: Blindering van patiënten en onderzoekers in een trial zorgt voor vergelijkbaarheid van baseline verschillen.

- ☐ juist
- ☐ onjuist

Vraag 16

Wordt een zogenaamd patiënt-controle onderzoek vaak gebruikt om inzicht te geven in de etiologie van een bepaalde ziekte?

- ☐ juist
- ☐ onjuist

Vraag 17

In een onderzoek naar de relatie tussen alcoholgebruik en het voorkomen van myocardinfarct, vergelijkt men een groep patiënten die in het afgelopen jaar een hartinfarct gehad heeft met een groep gezonde controles. Men vindt als odds ratio voor het drinken van meer dan twee glazen alcohol per dag 1,42 met een 95% betrouwbaarheidsinterval van 0,96-2,10. Wat concludeert u hieruit over de relatie tussen het drinken van meer dan twee glazen alcohol en het voorkomen van een myocardinfarct?

- ☐ Een statistisch significante positieve relatie. De p-waarde is kleiner dan 0,05.
- ☐ Een statistisch significante positieve relatie. De p-waarde is groter dan 0,05.
- ☐ Geen statistisch significante positieve relatie. De p-waarde is kleiner dan 0,05.
- ☐ Geen statistisch significante positieve relatie. De p-waarde is groter dan 0,05.

Vraag 18

In een studie werd gekeken naar het beschermend effect van ACE-inhibitors op de achteruitgang van nierfunctie bij patiënten met diabetes. Op een diabetespoli werd een random steekproef genomen van diabeten waarbij achteruitgang van de nierfunctie gedurende twee jaar was vastgesteld. Zowel bij de mensen met een relatief snelle achteruitgang van de nierfunctie als bij mensen met een relatief langzame achteruitgang van de nierfunctie werd gekeken naar de voorgeschreven medicatie (ACE-inhibitors) bij de start van die twee jaar. Wat voor onderzoeksopzet is dit?

- ☐ Patiënt-controle-onderzoek
- ☐ Prospectief follow-up onderzoek
- ☐ Randomised controlled trial
- ☐ Retrospectief follow-up onderzoek

Vraag 19

In een cohortonderzoek wordt gekeken of er een relatie bestaat tussen roken en het ontwikkelen van een bepaalde ziekte. De resultaten staan hieronder.

	Ziekte	
	Aanwezig	**Afwezig**
Roker	240	760
Niet-roker	260	1740

Bereken het relatieve risico op het ontwikkelen van de ziekte voor een roker ten opzichte van een niet-roker.

Risicocategorie (score)	**Patiënten***	**DVT aanwezig†**	**DVT afwezig‡**
Zeer laag (0-3)	23	0,7	99,3
Laag (4-6)	5	4,5	95,5
Matig (7-9)	51	21,7	78,3
Hoog (10-13)	21	51,3	48,7

Vraag 20

Bekijk de weergegeven tabel hieronder.

Prevalentie van DVT in 4 risicocategorieën bij gebruik van de diagnostische regel met D-dimeertest (%)

Stelling: Het gebruik van onderstaande risicoscore bij DVT is vooral geschikt om DVT aan te tonen.

- ☐ juist
- ☐ onjuist

Vraag 21

In een onderzoek naar de toegevoegde waarde van de bezinkingssnelheid (BSE) bij verdenking op pneumonie in de huisartspraktijk blijkt dat de oppervlakte onder de ROC curve van een diagnostisch model met gegevens uit anamnese en lichamelijk onderzoek en BSE 0,81 is. Een model met dezelfde gegevens uit anamnese en lichamelijk onderzoek, maar met C-reactief proteïne (CRP) in plaats van BSE is 0,90. Welke test voegt het meeste toe aan anamnese en lichamelijk onderzoek bij verdenking op pneumonie?

- ☐ BSE
- ☐ CRP
- ☐ Ze voegen allebei evenveel toe
- ☐ Ze voegen allebei niets toe

Vraag 22 tm 25

Hieronder staan 4 klinische vragen. Geef per vraag aan wat het domein is, d.w.z. voor welke patiënten/mensen zijn de resultaten van toepassing. Doe ditzelfde voor determinant en uitkomst. Belangrijk: niet bij elke klinische vraag is elk onderdeel aanwezig.

Vraag 22

Wat is het effect van antibiotica in vergelijking met expectatief beleid op de totale ziekteduur bij kinderen met otitis media acuta?

Vraag 23

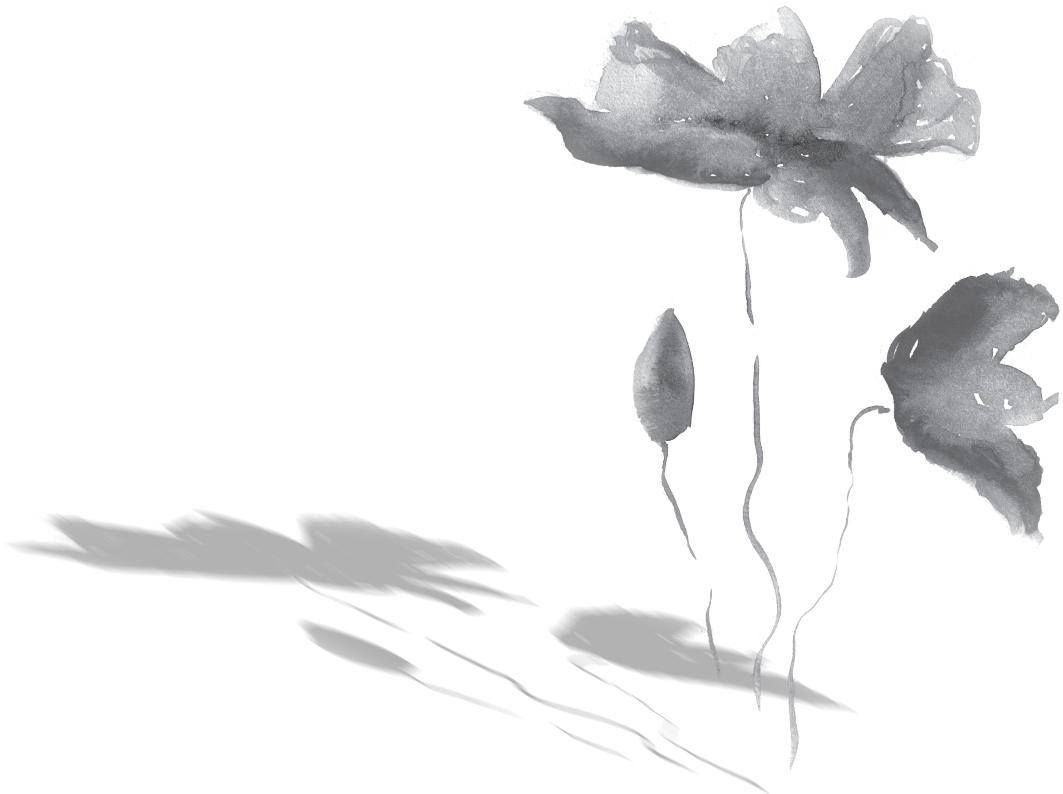
Kan een mazeleninfectie een hogere incidentie van 'inflammatory bowel disease' bij kinderen veroorzaken?

Vraag 24

Bestaat er een causaal verband tussen refluxziekte en het optreden van een chronische hoest?

Vraag 25

Wat is het effect van verrichten van een ECG bij patiënten met pijn op de borst in de huisartspraktijk op de verwijsfrequentie?



CHAPTER 7

Development and validation of a new instrument measuring guideline adherence in clinical practice

Based on: Development and validation of a new instrument measuring guideline adherence in clinical practice. Marlous F Kortekaas; Marie-Louise E L Bartelink; Geert J M G van der Heijden; Arno W Hoes; Niek J de Wit. Family Practice 2016; doi: 10.1093/fampra/cmw063

Kortekaas MF, Bartelink MEL,
van der Heijden GJMG, Hoes AW, de Wit NJ

"De meeste vrijheid geniet je binnen een structuur."
Marlous Kortekaas

ABSTRACT

Background

Education in evidence-based medicine (EBM) is an important part of the postgraduate training of general practitioners (GPs). Evaluation of its effect on EBM behaviour in daily clinical practice is difficult and instruments are scarce. Working in accordance with guidelines is considered as one of the key indicators of EBM behaviour.

Objective

To develop and validate an instrument assessing guideline adherence of GP trainees in clinical practice.

Methods

We developed an instrument that assesses guideline adherence, taking conscious deviation into account. The instrument assesses guideline adherence on 59 different management decisions (diagnosis N=17, therapy N=20, referral N=22) for 23 conditions as described in 27 different clinical practice guidelines. We validated this instrument using performance data as collected by third year GP trainees on three important properties; validity, reliability, and feasibility.

Results

Performance data were collected by 76 GP trainees on 12,106 patient consultations with 12,587 different reasons for encounter. Overall, guideline adherence was 82% (95%CI 77-88%). The significant correlation with the national GP knowledge test (r 0.33, p 0.004) showed the instrument to be a valid instrument. Interrater reliabilities (ICC) varied between moderate and excellent (0.64-1.00, p<0.001). The instrument proved feasible with coverage of 24% (N=3,082) of reasons for encounter presented to GP trainees and a mean and median time of one minute to score a patient consultation.

Conclusion

This instrument proved valid, reliable and feasible to assess guideline adherence among trainees in the clinical primary care setting.

INTRODUCTION

To date, evidence-based medicine (EBM) is well-integrated in medical curricula and in postgraduate training programs.(1) EBM enables healthcare professionals to make rational medical decisions by integrating the best available evidence with clinical expertise and patient values.(2) EBM is considered one of the core competencies in the postgraduate general practitioner (GP) training (3) and is integrated in this training in five steps following the Sicily Statement recommendations; ask, acquire, appraise, apply and assess.(4) The GP specialty training in the Netherlands takes three years, and combines 4 days training in clinical practice with one day theoretical education at the GP training institute at the University. The main goal is to teach GP trainees to use EBM into their daily practice. Therefore, it is important to be able to evaluate effects of EBM training in clinical practice.(4) There is, however, no agreement on the optimal way to assess EBM behaviour.(5) It is suggested that it is best assessed by monitoring clinical management decisions(6), but instruments for this are scarce.(5,7)

The clinical content of the postgraduate GP training in the Netherlands is based on the clinical practice guidelines (CPGs) for GPs from the Dutch College of General Practice.(8) Presently there are more than 100 evidence-based CPGs that are the result of a structured process based on collection and assessment of research evidence, and expert opinion.(8) Working in accordance with these guidelines is considered as one of the key indicators of practising EBM, and monitoring guideline adherence is a way to assess EBM behaviour during professional training.(9) We report on the development of an instrument to assess guideline adherence and its validation among a cohort of GP trainees at the UMC Utrecht.

METHODS

Instrument

Description of the instrument

This instrument assesses rational guideline adherence of GP trainees using patient consultation data. Rational adherence means clinical management in accordance with key recommendations in professional guidelines, thereby allowing motivated deviation in individual patients. The instrument assesses guideline adherence on 59 different management decisions (on diagnosis *N*=17, on therapy *N*=20, and on referral *N*=22) for 23 prevalent conditions, as described in 27 different CPGs.

Development process

We started with selecting quality indicators from an existing quality assessment instrument in primary care that is used to monitor guideline adherence among GPs. (10,11) We selected quality indicators that measure whether medical decisions adhered to the guideline recommendations. For pragmatic reasons we focused on those medical

decisions that do not require extensive knowledge of medical history or prior medical decisions.(12) To increase feasibility we selected quality indicators for disorders frequently presented to GP trainees.(13) This resulted in 19 quality indicators for 17 disorders (24% of all quality indicators in the existing quality assessment instrument). Based on the outcome of expert sessions (discussions with the academic GPs that were involved in the development of the original guideline) we adjusted the original quality indicators to be able to assess management decisions in individual patient consultations. An example of such an adjustment is shown in table 1. Pilot testing on anonymised patient data collected by GP trainees demonstrated that the 19 adjusted quality indicators assessed decisions in clinical management of approximately 20% of all patient consultations. Because we considered 20% low, we decided to extend the number of indicators in the instrument. Since in the first phase each quality indicator addressed only one management decision (on diagnosis, therapy or referral) we extended the instrument with assessment of additional management decisions from the same guidelines. Moreover, we expanded the number of guidelines involved, and designed new assessment items on management decisions for other disorders that were frequently presented to GP trainees. The final set of assessment items on management decisions was again extensively discussed with academic GP experts in one-to-one sessions.

Scoring guideline adherence

Guideline adherence was scored by two independent researchers (YS and LW) for every decision in clinical management (diagnosis, therapy, and referral) on a three-point-scale; (-1) not in accordance with the guideline and no reason mentioned to deviate from it, (0) debatable if in accordance with the guideline due to insufficient or contradicting information, and (+1) in accordance with the guideline or not in accordance but with rational motivation, for example in accordance to patient's wish or clinical expertise. In case of disagreements, a third researcher (JvD) made a final decision.

Validation process

Data collection

We validated the instrument in data from the PINET study (Personalized INtegrated Evidence-based medicine teaching for Trainees in general practice); a prospective, cluster randomized controlled trial, in which third-year GP trainees from Utrecht University were allocated to either an integrated EBM training program or a stand-alone EBM training program. Demographics and baseline characteristics of all participants were collected. We used data from two consecutive measurements, at the beginning and end of the third year of the postgraduate GP training programme. GP trainees collected the following data from the electronic medical records (EMR) in the GP trainee practice in logs: gender, age, reason for encounter, medical history, results of physical examination, diagnosis (ICPC code(14)), and treatment or referral. If a patient presented multiple reasons for encounter during one consultation, multiple ICPC codes were allocated.

We used ICPC codes to select patient consultations regarding disorders described in the instrument (table 2). We excluded the following consultations from assessment; 1) a follow-up consultation, as prior management decisions affect the GP trainee's decision, 2) a patient consultation that fits in multiple guidelines addressing a single reason for encounter, (for example; common cold is addressed by the guidelines on "acute cough", "rhinosinusitis", and "acute sore throat"), 3) missing data on age or gender, when relevant (e.g. in a patient consultation about asthma where the management decision depends on the age of the patient)(15,16), and 4) a patient consultation in which the aim of the consultation does not match the selected assessment items (e.g. a patient with diabetes mellitus who wants to talk about the social implications of the disease).

Psychometric properties

Validity

Validity, defined as the degree to which an instrument measures the construct(s) it purports to measure.(17) For this we evaluated content-, and construct validity. Content validity has been defined as the degree to which the content of the instrument adequately reflects the construct to be measured. We first ascertained this by taking quality indicators from an existing, validated quality assessment instrument.(10,11) Next, we discussed the different subjects of the instrument with nine academic GPs separately, each of them selected because they were expert in the specific subject of the guideline, and involved in its development. Adjustments and extensions of the instrument on their field of interest were extensively discussed with them in one-to-one sessions. Subsequently, they were asked for their final approval Construct validity has been defined as the degree to which the scores of the instrument are consistent with hypotheses (for instance with regard to relationships to scores of other instruments) based on the assumption that the instrument validly measures the construct to be measured.(17) We compared GP trainees' guideline adherence at baseline (mean) to the mean baseline score on the national GP knowledge test (LHK), using the Spearman rank correlation coefficient. The LHK is a test measuring knowledge of the professional guidelines. It consists of 120 multiple-choice questions about the national GP CPGs, and is used to monitor clinical knowledge development of all GP trainees twice a year.(18) For each GP trainee guideline adherence was assessed as the number of times a GP trainee adhered to a guideline (i.e. scored +1) as a proportion of all reasons for encounter in the patient consultations assessed with the instrument as seen by that trainee.

Reliability

Reliability has been defined as the degree to produce stable and consistent results. (17) For reliability we measured the extent to which the two independent raters (both last-year medical students) produced the same results on the same data (interrater reliability).(17) We calculated the absolute intraclass correlation coefficient ($ICC_{absolute}$ (2,1)) for three decisions.(19,20) First for the inclusion (or exclusion) of a patient consultation after the first selection, that was based on ICPC codes. Second was

the choice of the assessment item to score a management decision, since for some disorders the choice of assessment item depends on age and the (exact) diagnosis, as interpreted by the rater. For example, there are two asthma CPGs (adults vs. children), and three for knee complaints (traumatic vs. non-traumatic, adults vs. children).(15,16,21–23) Third was the allocated guideline adherence score to a management decision on diagnosis, therapy, or referral. ICC values were interpreted as: excellent reliability ≥0.8, moderate reliability is 0.61-79, and questionable reliability <0.60.(19)

Feasibility

Feasibility was defined as the viability to assess guideline adherence in a substantial number of patient consultations in a short amount of time per consultation. For feasibility we first calculated the coverage of the instrument, defined as the number of reasons for encounter in patient consultations presented to GP trainees that were scored, divided by the total number of reasons for encounter in these consultations. In addition, we assessed the time needed to assess patient consultations (in minutes).

Data analyses

Statistical analyses were performed using Statistical Package for the Social Sciences version 20.0 (SPSS Inc., Chicago, IL, USA), and SAS version 9.2 (SAS Institute Inc., Cary, North Carolina). A *p*-value of ≤ 0.05 was considered to be statistically significant.

Ethical approval

This study design was assessed by the UMC Utrecht Ethics Committee, and regarded as non-eligible for full informed consent. However, we obtained informed consent from the GP trainees for the use of the data from the logs. Patient data were reported anonymously.

RESULTS

Validation process

Data collection

We used data from 12,106 patient consultations, with 12,587 different reasons for encounter, performed by 76 GP trainees at two measurements in the third year of the postgraduate GP training programme. In line with the gender distribution of the GP trainees in the Netherlands trainees in our study were predominantly female (72%), and the median age was 31 years (IQR 5).(24,25) Most (95%) trainees did a hospital internship before starting the GP vocational training program for a median time period of 20 months (IQR 19). Two (3%) trainees had a PhD. Two third (68%) of the trainees worked fulltime during the GP training. The most frequent reasons for encounter presented to the GP trainees were skin problems (ICPC category S, 19%), musculoskeletal disorders (L, 18%), and respiratory complaints (R, 17%). Based on predefined ICPC codes (table 2) we selected 4,203 patient consultations (33% of total)

Table 1. An example of an adjustment of an existing quality indicator (11) to an assessment item as included in the developed instrument used to assess guideline adherence.

Original quality indicator				
Acute otitis media (AOM) in children older than two years (36)	Numerator: Children >2yrs old with AOM who did not receive antibiotics (AB). Denominator: All children >2yrs with AOM.	Percentage of children > 2yrs with AOM that did not receive AB.		
Adjusted quality indicator (assessment item)				
Acute otitis media (AOM) in children (36)	No AB, unless indicated^. When AB indicated, according to the schedule* as provided in the GP guideline.(32)	-1	AB	AB prescribed when not indicated^ or not according to schedule* in GP guideline, and no reason mentioned (e.g. parent's wish, allergy). Quinolones and/or lidocaine prescribed as eardrops.
			No AB	No AB* prescribed when indicated^. No AB* prescribed despite otorrhea at first presentation, and no advice given to come back when otorrhea lasts for more than one week.
		0	AB	AB prescribed as eardrops (not being quinolones or lidocaine).
			?	Unclear what sort of AB was prescribed. Unclear what age the child had.
			No AB	No AB* prescribed while indicated^, but no pain or illness at the same time.
		1	AB	AB prescribed when indicated^ , and according to schedule* in GP guideline, or deviated from guideline with reason (e.g. parent's wish, allergy)..
			No AB	No AB* prescribed when not indicated^.

^AB indicated according to GP guideline: 1) >3days of complaints with fever, 2) <3 days of complaints, but very ill or quickly becoming more ill, 3) patient in risk group, as defined in GP guideline, 4) Otorrhea at first presentation or for more than one week, 5) AOM in both ears and child <2yrs old.
*AB schedule according to GP guideline: 1) Amoxicillin (7 days), 2) Azithromycin (3 days) or co-trimoxazole (5-7 days) in case of contra-indications for amoxicillin, 3) Augmentin in case of otorrhea in ill children with ear tubes.

for assessment. We then excluded 1121 patient consultations (9% of total) based on the exclusion criteria, resulting in a final selection of 3,082 patient consultations (24% of total) with 8,083 different management decisions; 2,416 diagnostic (19% of total), 2,585 therapeutic (21% of total), and 3,082 referral decisions (24% of total).

Psychometric properties

Validity

All nine academic GP experts considered the assessment items on their subject in the instrument to be relevant for daily clinical practice and an adequate reflection of the construct to be measured. The evaluation of construct validity showed that the mean baseline score of the 76 GP trainees on guideline adherence was significantly correlated to their baseline result on the LHK (*r* 0.33, *p* 0.004).

Reliability

The interrater reliability on inclusion of a patient consultation for assessment (*N*=3,010) and on the choice of assessment item to score a management decision (*N*=2,939) was excellent with, an ICC of 0.80 (p<0,001, 95% CI 0.79-0.81) and 1.00 (p<0,001, 95% CI 1.00-1.00), respectively. Disagreement on inclusion occurred in 8% of patient consultations. Disagreement on the choice of assessment item for a management decision was rare (*N*=71, 2% of total), and mainly occurred when the reason for encounter was related to knee problems (*N*=53, 75% of disagreements) where the assessment items in the instrument are based on three different guidelines.(20–22) The interrater reliability for the allocated guideline adherence scores (-1, 0 or +1) to all assessed management decisions (*N*=8083) was excellent with an ICC of 1.00 (p< 0.001, 95% CI 1.00-1.00). For diagnostic (*N*=2,416), therapeutic (*N*=2,585), and referral (*N*=3082) decisions the interrater reliabilities were moderate with ICCs of 0.79 (*p*<0.001, 95% CI 0.78-0.81), 0.79 (*p*<0.001, 95%CI 0.77-0.80), and 0.64 (*p*<0.001, 95%CI 0.62-0.66), respectively.

Feasibility

The instrument we developed covered 24% (*N*=3,082) of all reasons for encounter (*N*=12,587), presented during the consultations. The mean and median time for both raters to assess patient consultations was around 1 minute; (1.26 (SD 0.5) and 1.04 (SD 0.2) minutes, range 1-7 minutes.

Guideline adherence scores

In table 3 the distribution of guideline adherence scores is shown. For the vast majority of the decisions in clinical management we assessed (*N*=8,083, 79%), GP trainees adhered to the guidelines, especially for the diagnostic decisions (*N*=1,930, 80%) and decisions on referral (*N*=2,759, 90%). In 8% of patient consultations (N=250) trainees appropriately motivated why they deviated from the recommendations in the guideline, with as primary reason patient's preferences (51%). For only a minority of

Table 2. Summary of clinical disorders, national general practice guidelines, and International Classification of Primary Care (ICPC) codes included in the instrument.

Clinical disorders	National general practice guidelines	ICPC codes (14)
Deep vein thrombosis	Deep vein thrombosis (37)	K94
Atrial flutter	Atrial flutter (38)	K78, K04, K05
Cardiac failure	Cardiac failure (39)	K77
Acute cough	Acute cough (40)	R71, R74, R77, R78, R81
Acute sore throat	Acute sore throat (41)	R21, R22, R72, R76
Asthma	Asthma in adults (15) Asthma in children (16)	R96
COPD	COPD (42)	R95
Allergic rhinitis/hay fever	Allergic and not-allergic rhinitis (43)	R97, R07
Rhino sinusitis	Rhino sinusitis (44)	R07, R09, R75
Stomach complaints	Stomach complaints (45)	D02, D03, D09, D10, D74, D84, D85, D86, D87
Cystitis/ pyelonephritis	Urinary tract infections (46)	U70, U71
Bacterial skin infections	Bacterial skin infections (47)	S09, S10, S11, S76, S84
Dermatomycosis	Dermatomycosis (48)	S74, S75
Acne	Acne (49)	S96
Shoulder complaints	Shoulder complaints (50)	L76, L92, L99
Backpain	Non-specific lower back pain (51) Lumbar radicular syndrome (52)	L02, L03, L08, L71, L76, L84, L86, L99
Knee complaints	Non-traumatic knee complaints in adults (21) Non-traumatic knee complaints in children (23) Traumatic knee complaints (24)	L15, L70, L71, L73, L75, L76, L78, L80, L90, L94, L96, L97, L99
Ganglion	Hand- en wrist complaints (53)	L87
Ankle complaints	Ankle ligament injury (54)	L16, L73, L77
Acute otitis media in children	Acute otitis media in children (36)	H71
Otitis media with effusion	Otitis media with effusion (55)	H72
Otitis externa	Otitis externa (56)	H70
Conjunctivitis and blepharoconjunctivitis	The red eye (57)	F70, F71, F72

the management decisions (*N*=443, 5%) GP trainees did not adhere to the guidelines, where for the remaining 16% it was debatable if GP trainees adhered to the guidelines due to insufficient or contradicting information. Information was missing in 24% of these patient consultations, mainly regarding the medical history (21%), or background information (e.g. comorbidity, allergies, etc.) on the patient (19%). Ten GP trainees (7% of 136) adhered to the guidelines on all diagnostic decisions (with a median of 7 patient consultations per trainee), and 38 trainees (28%) on all referral decisions (with a median of 16 patient consultations per trainee). In 64% of the therapeutic decisions (*N*=2,585) GP trainees adhered to the guidelines. In total, 58% of the 3,082 patient consultations was completely in accordance with the guidelines.

Table 3. Guideline adherence scores in management decisions of 76 third-year GP trainees.

Score	Diagnostic decisions	Therapeutic decisions	Referral decisions	All decisions
-1	108 (4)	311 (12)	240 (1)	443 (5)
0	378 (16)	64 (24)	299 (10)	1,288 (16)
1	1,930 (80)	1,663 (64)	2,759 (90)	6,352 (79)
Total	**2,416 (100)**	**2,585 (100)**	**3,082 (100)**	**8,083 (100)**

(-1) not in accordance with the guideline and no reason mentioned
(0) debatable if it is in accordance with the guideline due to insufficient or contradicting information
(+1) in accordance with the guideline or a conscious deviation from it

DISCUSSION

Summary

To our knowledge this is the first instrument assessing guideline adherence of GP trainees in clinical practice, also allowing motivated deviation in individual patients. The instrument enables assessment of three relevant aspects of clinical management decisions (diagnosis, therapy, and referral) for a wide variety of disorders, covering one in four GP trainees' patient consultations. The instrument was demonstrated to be valid, reliable and feasible to use, and facilitates the monitoring of guideline adherence as one of the key indicators of EBM behaviour at the actual point of care during the GP postgraduate training program.

Strengths and limitations

A major strength is the ability of the instrument to use data from patient consultations of the trainee for a wide range of disorders. The instrument is based on guidelines that are widely accepted and applied in daily practice.(8) We validated the instrument in a large number of patient consultations, assessing the recommended properties (validity, reliability, and feasibility).(5,7) We do realize that optimal EBM behaviour in primary care extends beyond guideline adherence alone. Personalized care means taking the best decisions in the perspective of the patient. However, we think that guideline adherence is an important aspect of EBM behaviour, as guidelines are created as a result of a combination of research evidence and clinical expertise, and, more importantly, because we also allowed for deviations from guideline recommendations when appropriately motivated, for example based on patient preferences.(2,8,26) The method we used to collect patient data for testing the instrument may have led to incomplete data, as we were highly dependent on the way GP trainees recorded their patient consultations. Previous studies show that only one-third of actions undertaken in a patient consultation are recorded with large differences between different parts of the consultations.(27) Misclassification of allocated guideline adherence scores may have occurred when trainees did not record reasons to deviate from a guideline. However, with the growing policy needs and demands with respect to political, juridical and insurance aspects in

health care, it is expected that patient records will become more comprehensive, (partly) resolving this limitation.(28) Moreover, as only one fifth of all the decisions was scored as "debatable whether in accordance with the guideline due to insufficient or contradicting information" (score 0), we think dependency on missing data was not an important issue in our instrument. For this reason we expect that our instrument could be used among GPs as well, even though they (as more experienced doctors) are expected to record less and therefore score worse.(29) It can be argued that the instrument covers only 24% of all reasons for encounter. Prior research from 1996 (when only 60 national guidelines were available) showed that at that time 27% of all diseases encountered in general practice were covered by the guidelines.(30) Extrapolated to the present number of 92 guidelines, the instrument's coverage could be regarded as relatively good, as it addresses around 40% of all reasons for encounters. Finally, we do realize that the national GP knowledge test (LHK) is not the optimal indicator of professional development, as it basically measures theoretical knowledge of professional guidelines. (18) However, as no better alternatives are available and as it is widely used in GP training, we think it provides an acceptable benchmark of progress during GP training.

Comparison with existing literature

Previously developed instruments that assess EBM behaviour do not assess guideline adherence, and most of them use indirect methods (i.e. through questionnaires, interviews, and focus groups) and not clinical practice data.(5) Adherence to guidelines has been assessed before in studies among GPs, but has not been used for this purpose before.(10,31,32) Compared to these previous studies, where motivated deviation in individual patients was not taken into account, guideline adherence seems comparable, with a slightly better performance among our GP trainees.(10,31,32) The main reason trainees mentioned to deviate from a guideline was the patient's preference, which aligns perfectly with the definition of EBM, as also argued by Greenhalgh et.al.(33)

Implications for future

The instrument we developed to measure guideline adherence in primary care practice is valid, reliable, and feasible. It could be used to monitor clinical performance during professional training.(9) Moreover, trainees would be able to reflect on EBM practice behaviour by evaluating their patient consultations using the instrument and discussing the results. When patient data can be derived from EMRs automatically, more (background) information will be available (such as patient characteristics, co-morbidity, etc.) enabling optimization of the assessment process.(34) As the use of guidelines is well-established and widely accepted among different specialties and in different countries nowadays(26,35), the concept we used can be applied to other health care settings as well. Our instrument could serve as an example to create an instrument measuring guideline adherence as an indicator of EBM behaviour in clinical practice in a valid, reliable, and feasible way.

REFERENCES

1 Ilic D, Nordin R Bin, Glasziou P, Tilson JK, Villanueva E. Development and validation of the ACE tool: assessing medical trainees' competency in evidence based medicine. BMC Med Educ [Internet]. 2014 Jan [cited 2014 Aug 19];14:114.

2 Sackett DL, Rosenberg WMC, Gray J a M, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. 1996. BMJ. 1996 Feb;312(71).

3 Frank JR. The CanMEDS 2005 physician competency framework. Better standards. Better physicians. Better care. Ottawa R Coll Physicians Surg Canada. 2005;

4 Dawes M, Summerskill W, Glasziou P, Cartabellotta A, Martin J, Hopayian K, et al. Sicily statement on evidence-based practice. BMC Med Educ. 2005 Jan 5;5(1):1.

5 Oude Rengerink K, Zwolsman SE, Ubbink DT, Mol BWJ, van Dijk N, Vermeulen H. Tools to assess evidence-based practice behaviour among healthcare professionals. Evid Based Med. 2013 Aug;18(4):129–38.

6 Tilson JK, Kaplan SL, Harris JL, Hutchinson A, Ilic D, Niederman R, et al. Sicily statement on classification and development of evidence-based practice learning assessment tools. BMC Med Educ [Internet]. BioMed Central Ltd; 2011;11(1):78.

7 Shaneyfelt T, Baum KD, Bell D, Feldstein D, Houston TK, Kaatz S, et al. Instruments for evaluating education in evidence-based practice: a systematic review. JAMA. 2006 Sep 6;296(9):1116–27.

8 nhg.org [Internet]. Available from: https://www.nhg.org

9 De Belvis a. G, Pelone F, Biasco a., Ricciardi W, Volpe M. Can primary care professionals' adherence to Evidence Based Medicine tools improve quality of care in Type 2 diabetes mellitus? A systematic review. Diabetes Res Clin Pract. 2009;85(2):119–31.

10 Braspenning J, Pijnenborg L, in 't Veld C, Grol R. Werken aan kwaliteit in de huisartsenpraktijk. Indicatoren gebaseerd op de NHG-standaarden. First. Houten: Bohn Stafleu van Loghum; 2005. 145 p.

11 Braspenning J, Schellevis F, Grol R. Tweede Nationale Studie naar ziekten en verrichtingen in de huisartspraktijk. Kwaliteit huisartsenzorg belicht. NIVEL/WOK, editor. Utrecht/Nijmegen; 2004.

12 Donabedian A. An Introduction to Quality Assurance in Health Care. New York: Oxford University Press; 2002.

13 Jong J De, Mohrs J, Waard MW. Opening the black box : the patient mix of GP trainees. 2011;(October):650–7.

14 Lamberts H WM. ICPC. International Classification of Primary Care. Oxford: Oxford University Press; 1987.

15 Bindels PJE, Van der Wouden JC PB, Brand PLP, Salomé PL, Van Hensbergen W V, Hasselt PA, Steenkamer TA GM. NHG-Standaard Astma bij kinderen. Huisarts Wet. 2006;49(11):557–72.

16 Geijer RMM, Chavannes NH, Muris JWM S, APE, Schermer T, Smeele IJM, Thoonen B V der, Molen T, Van Schayck CP, Van Weel C K, BGM G. NHG-Standaard Astma bij volwassenen. Huisarts Wet. 2007;50(11):537–51.

17 Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol [Internet]. 2010 Jul;63(7):737–45.

18 Van Leeuwen YD, Pollemans MC, Mol SS, Eekhof JA, Grol R, Drop MJ. The Dutch knowledge test for general practice: issues of validity. Eur J Gen Pract 1995;1:113-7.

19 Richman J, Makrides L, Prince B. Research methodology and applied statistics. Physiother Can. 1980;32(4):253–7.

20 Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34–42.

21 Belo JN, Bierma-Zeinstra SMA, Raaijmakers AJ, Van der Wissel F OW. NHG-Standaard Niet-traumatische knieproblemen bij volwassenen (Eerste herziening). Huisarts Wet. 2008;51(5):229–40.

22 Belo JN, Berg HF, Klein Ikkink AJ, Wildervanck-Dekker CMJ, Smorenburg HAAJ DL. NHG-Standaard Traumatische knieproblemen (Eerste herziening). Huisarts Wet. 2010;54(3):147–58.

23 Breedveldt Boer HP, Klaassen WRC, Spinnewijn WEM. Heinen N, Burggraaff HB, Derks CJT LM. NHG-Standaard Niet-traumatische knieproblemen bij kinderen en adolescenten(Eerste herziening). Huisarts Wet. 2009;52(7):332–41.

24 Vermeulen MI, Kuyvenhoven MM, Zuitho NPA, van der Graaf Y, Damoiseaux RAMJ. Dutch postgraduate GP selection procedure; reliability of interview assessments. BMC Fam Pract 2013,14:43 doi: 10.1186/1471-2296-14-43.

25 Jong J De, Mohrs J, Waard MW. Opening the black box : the patient mix of GP trainees. 2011;(October):650–7.

26 Ansari S, Rashidian A. Guidelines for Guidelines: Are They Up to the Task? A Comparative Assessment of Clinical Practice Guideline Development Handbooks. PLoS One. 2012;7(11).

27 Rethans JJ, Martin E, Metsemakers J. To what extent do clinical notes by general practitioners reflect actual medical performance? A study using simulated patients. Br J Gen Pract. 1994;44(381):153–6.

28 Koerselman G, Korzec A. Voorstel voor een checklist bij het afwijken van richtlijnen. NTvG. 2008;152:1757–9.

29 Glasziou P, Moynihan R, Richards T. Too much medicine ; too little care. Br Med J. 2013;347(2 July):1–2.

30 Wiersma TJ, Rutten GEHM. Standaarden en epidemiologie anno 1996. Thomas S, Geijer RMM, Laan van der JR, Wiersma TJ. NHG-standaarden voor de huisarts II. Utrecht: Bunge, 1996.

31 Van den Berg MJ, de Bakker DH, Spreeuwenberg P, Westert GP, Braspenning JCC, van der Zee J, et al. Labour intensity of guidelines may have a greater effect on adherence than GPs' workload. BMC Fam Pract. 2009;10:74.

32 Willekens M, Giesen P, Plat E, Mokkink H, Burgers J GR. Quality of after-hours primary care in The Netherlands: adherence to national guidelines. BMJ Qual Saf. 2011;20(3):223–7.

33 Greenhalgh T, Howick J, Maskrey N. Evidence based medicine: a movement in crisis? Bmj [Internet]. 2014;348(jun13 4):g3725–g3725.

34 Van Berkestijn LGM, Kastein MR, Lodder A, De Melker R a., Bartelink ML. How do we compare with our colleagues? Quality of general practitioner performance in consultations for non-acute abdominal complaints. Int J Qual Heal Care. 1999;11(6):475–86.

35 Wolf JS, Hubbard H, Faraday MM, Forrest JB. Clinical practice guidelines to inform evidence-based clinical practice. World J Urol. 2011;29(3):303–9.

36 Damoiseaux RAMJ, Van Balen FAM, Leenheer WAM KB. NHG-Standaard Otitis media acuta (Tweede herziening). Huisarts Wet. 2006;49(12):615–21.

37 Oudega R, Van Weert H, Stoffers HEJH, Sival PPE, Schure RI, Delemarre J, Eizenga WH. NHG-Standaard NHG-Standaard Diepe veneuze trombose en longembolie. Huisarts Wet 2008;51(1):24-37.

38 Boode BSP, Heeringa J, Konings KTS, Rutten FH, Van den Berg PJ, Zwietering PJ, Van Lieshout J, Opstelten W. NHG-Standaard Atriumfibrilleren (Eerste herziening). Huisarts Wet 2009:52(13):646-63.

39 Hoes AW, Voors AA, Rutten FH, Van Lieshout J, Janssen PGH, Walma EP. NHG-Standaard Hartfalen (Tweede herziening). Huisarts Wet 2010:53(7):368-89.

40 Verheij ThJM, Hopstaken RM, Prins JM, Salomé PhL, Bindels PJ, Ponsioen BP†, Sachs APE, Thiadens HA, Verlee E. NHG-Standaard Acuut hoesten (Eerste herziening). Huisarts Wet 2011:54(2):68-92.

41 Zwart S, Dagnelie CF, Van Staaij BK, Balder FA, Boukes FS, Starreveld JS. NHG-Standaard Acute keelpijn (Tweede herziening). Huisarts Wet 2007;50(2):59-68.

42 Smeele IJM, Van Weel C, Van Schayck CP, Van der Molen T, Thoonen B, Schermer T, Sachs APE, Muris JWM, Chavannes NH, Kolnaar BGM, Grol MH, Geijer RMM. NHG-Standaard COPD (tweede herziening). Huisarts Wet 2007;50(8):362-79.

43 Sachs APE, Berger MY, Lucassen PLBJ, Van der Wal J, Van Balen JAM, Verduijn MM. NHG-Standaard Allergische en niet-allergische rhinitis (Eerste herziening). Huisarts Wet 2006;49(5);254-65.

44 De Sutter A, Burgers JS, De Bock GH, Dagnelie CF, Labots-Vogelesang SM, Van der Laan JR, et al. NHG-Standaard Rhinosinusitis (tweede herziening). Huisarts Wet 2005:48:615-26).

45 Numans ME, De Wit NJ, Dirven JAM, Heemstra-Borst CG, Hurenkamp GJB, Scheele ME, Burgers JS, Geijer RMM, De Jongh E. NHG-Standaard Maagklachten(Derde herziening). Huisarts Wet 2013;56:26-35.

46 M A K Van Haaren, S H Visser, S Van Vliet, E A Timmermans, R Yadava, E S Geerlings, G Ter Riet, B Van Pinxteren. NHG-Standaard Urineweginfectie (Tweede herziening). Huisarts Wet 2005(8):341-52.

47 Wielink G, Koning S, Oosterhout RM, Wetzels R, Nijman FC, Draijer LW. NHG-Standaard Bacteriële huidinfecties (Eerste herziening). Huisarts Wet 2007;50(9): 426-44.

48 Van Puijenbroek EP, Duyvendak RJP, De Kock CA, Krol SJ, Jaspar AHJ, Loogman MCM. NHG-Standaard Dermatomycosen (Eerste herziening). Huisarts Wet 2008;51(2):76-84.

49 Smeets JGE, Grooten SJJ, Bruinsma M, Jaspar AHJ, Kertzman MGM. NHG-Standaard Acne (Tweede herziening). Huisarts Wet 2007:50(6):259-68.

50 Winters JC, Van der Windt DAWM, Spinnewijn WEM, De Jongh AC, Van der Heijden GJMG, Buis PAJ, Boeke AJP, Feleus A, Geraets JJXR. NHG-Standaard Schouderklachten (Tweede herziening). Huisarts Wet 2008:51(11):555-565

51 Chavannes AW, Mens JMA, Koes BW, Lubbers WJ, Ostelo R, Spinnewijn WEM, Kolnaar BGM. NHG-Standaard Aspecifieke lagerugpijn (Eerste herziening). Huisarts Wet 2005;48(3):113-23.

52 Mens, J.M.A., Chavannes, A.W., Koes, B.W., Lubbers, W.J., Ostelo, R.W.J.G., Spinnewijn, W.E.M. & Kolnaar, B.G.M. (2005). NHG-standaard. Lumbosacraal radiculair syndroom. Eerste herziening. Huisarts Wet. 2005;48(4):171-8.

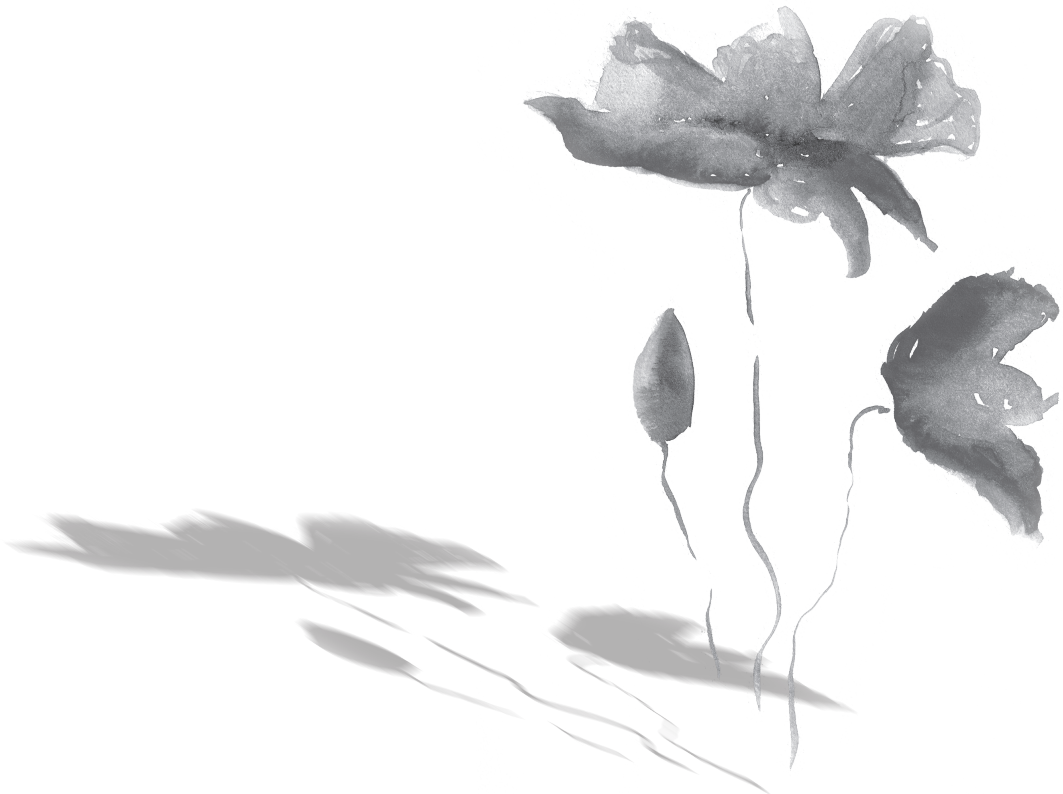
53 Peters-Veluthamaningal C, Willems W, Smeets JGE, Van der Windt DAWM, Spies MN, Strackee SD, Vos K, Wind LA, Geraets JJXR. NHG-Standaard Hand- en polsklachten. Huisarts Wet 2010:53(1):22-39.

54 Janneke Belo, Pieter Buis, Rogier van Rijn, Evelien Sentrop-Snijders, Sicco Steenhuisen, Clara Wilkens, Roeland Geijer, Masja Loogman. NHG-Standaard Enkelbandletsel (tweede herziening). Huisarts Wet 2012;55(8).

55 Van Balen FAM, Rovers MM, Eekhof JAH, Van Weert HCPM, Eizenga WH en Boomsma LJ. NHG-Standaard Otitis media met effusie (tweede herziening). In: Wiersma T, Goudswaard AN, redactie. NHG-Standaarden voor de huisarts. Houten: BSL, 2006;901-12.

56 Rooijackers-Lemmens E, Balen FAM van, Opstelten W, Wiersma Tj. NHG-standaard Otitis externa. Eerste herziening. In: Wiersma Tj, Goudswaard AN, redacteuren. NHG-standaarden voor de huisarts. Versie 2006. Houten: Bohn Stafleu van Loghum; 2006. p. 885-93.

57 Rietveld RP, Cleveringa JP, Blom GH, Baggen MEJM, Bink D, Oltheten JMT, Van der Weele GM. NHG-Standaard Het rode oog (Eerste herziening). Huisarts Wet 2006;49(2):78-91.



CHAPTER 8

Does integrated training in evidence-based medicine (EBM) in the general practice (GP) specialty training improve EBM behaviour in daily clinical practice? A cluster randomized controlled trial

Accepted for publication in BMJ Open. Does integrated training in evidence-based medicine (EBM) in the general practice (GP) specialty training improve EBM behaviour in daily clinical practice? A cluster randomized controlled trial.

Kortekaas MF, Bartelink MEL, Zuithoff NPA,
vd Heijden GJMG, de Wit NJ, Hoes AW

"Goed EBM-onderwijs moet tegen kritiek kunnen."
Marlous Kortekaas

ABSTRACT

Objectives

Evidence-based medicine (EBM) is an important element in the general practice (GP) specialty training. Studies show that integrating EBM training into clinical practice brings larger benefits than stand-alone modules. However, these studies have neither been performed in general practice nor assessed EBM behaviour of former trainees in daily clinical practice.

Setting

GP specialty training in the Netherlands.

Participants

All 82 third-year GP trainees who started their final third year in 2011 were approached for inclusion, of whom 79 (96%) participated; 39 in the intervention and 40 in the control group.

Intervention

Integrated EBM training, in which EBM is embedded closely within the clinical context by joint assignments for the trainee and supervisor in daily practice, and teaching sessions based on dilemmas from actual patient consultations.

Comparison

Stand-alone EBM training at the institute only.

Primary and secondary outcomes

Our primary outcome was EBM behaviour, assessed by measuring guideline adherence (incorporating rational, motivated deviation) and information-seeking behaviour. Our secondary outcomes were EBM attitude and EBM knowledge. Data were acquired using logbooks and questionnaires, respectively. Analyses were performed using mixed models.

Results

Logbook data were available from 76 (96%) of the participating trainees at baseline (7,614 consultations), 60 (76%) at the end of the third year (T1, 4,973 consultations), and 53 (67%) one year after graduation (T2, 3,307 consultations). We found no significant differences in outcomes between the two groups, with relative risks for guideline adherence varying between 0.96 and 0.99 (95%CI 0.86-1.11) at T1, and 0.99 and 1.10 (95%CI 0.92-1.25) at T2, and for information-seeking behaviour between 0.97 and 1.16 (95%CI 0.70-1.91), and 0.90 and 1.10 (95%CI 0.70-1.32), respectively.

Conclusion

Integrated EBM training compared to stand-alone EBM training does not improve EBM behaviour, attitude or knowledge of (future) GPs.

INTRODUCTION

Evidence-based medicine (EBM), defined as integrating clinical expertise, patient values and the best available clinical evidence in daily clinical practice is the cornerstone of today's medical practice.(1) General practice (GP) trainees in the Netherlands learn to work according to the principles of EBM during their three-year competency-based specialty training.(2)

EBM is taught according to the five steps as defined in the Sicily Statement; ask, acquire, appraise, apply and assess.(3) Despite its undisputed importance, evidence on the optimal method to teach EBM is lacking.(4) Traditionally, EBM in the GP specialty training in the Netherlands is trained in theoretical, stand-alone educational sessions. This differs from integrated EBM training, in which EBM is trained in a clinical context, and teaching sessions are based on recent patient consultations in the trainees' practice. A review from 2004 reported that EBM knowledge among hospital trainees improved with both stand-alone and integrated EBM training, while skills, attitude and behaviour improved only with integrated EBM training. The authors concluded that " teaching of evidence-based medicine should be moved from classrooms to clinical practice to achieve improvements in substantial outcomes".(5) However, the effectiveness of stand-alone and integrated EBM training was never directly compared in a (randomized) controlled trial, nor were the effects on EBM behaviour of former trainees assessed in daily clinical practice after graduation.(5,6) Moreover, it is unknown whether integrated EBM training also improves EBM behaviour in GP trainees; in primary care both the setting (one supervisor and one trainee) and the patient and disease spectrum differ substantially from the hospital environment.

We developed an integrated EBM training for GP trainees with the focus on the last two steps of EBM, i.e. apply and assess (table 1). We also trained the supervising GPs, stressing the important role these supervisors have in integrating EBM training into clinical practice.(7) We report the results of a cluster randomized controlled trial among third year GP trainees comparing the effects of this integrated EBM training program with a stand-alone EBM training program.

METHODS

Design

The PINET study (Personalized INtegrated Evidence-based medicine teaching for Trainees in general practice) was designed as a cluster randomized controlled trial among third (last) year GP trainees from six different groups (with 12 trainees each) starting at four consecutive time moments. Groups of trainees were randomly allocated to either an integrated (i.e. new) EBM training program (intervention group) or the regular stand-alone EBM training program (control group).

Study population

The study was performed within the GP specialty training of the University Medical Centre Utrecht (UMCU) in the Netherlands between March 2011 and December 2013; a 3-year specialization program for general practitioners. The 1st and 3rd year consist of in practice training under the supervision of an experienced GP, the second year consists of hospital rotations.(8) All 82 third-year GP trainees who started their final third year in 2011 were approached for inclusion.

Allocation and data collection

In total 6 trainee groups successively entered their 3rd trainee year in 2011 (two groups in March, one in June, two in September, and one in December). As trainee groups of two consecutive time moments (2 groups in March and 1 in June, and 2 in September and 1 in December) have joint educational programs, we clustered the intervention in time, and randomly allocated the first three groups to the intervention and the last three groups to the control arm. Data were collected at baseline (T0), at the end of the third year (T1), and one year after graduation (T2). Baseline characteristics such as age, sex, medical school, clinical experience, and score on the GP knowledge test, a national test that is used to monitor knowledge progress, were collected.(9) Trainees were also asked to self-assess their EBM attitude and EBM knowledge at baseline on a 5-point Likert scale ranging from 1 ("very poor") to 5 ("very good").

Intervention

The EBM training program in the intervention and control group is summarized in table 1. Essential skills, such as searching for evidence, critical appraisal of the literature for different research designs, and basic analytic skills are taught in accordance with the five steps of EBM training as described in the Sicily Statement, with the focus on clinical relevance as (future) GPs.(3) The main difference between the stand-alone and integrated EBM training program is the focus on the last two steps of EBM in the latter training program. The theoretical basis of the program was therefore adapted to emphasize the practical implication of research and to stress its clinical relevance. (10,11) In addition, we changed four components of the (regular) stand-alone EBM training program and incorporated two new elements in the program to better link EBM theory to clinical practice (see table 1 for details). Finally, their GP supervisors in primary care practice were involved in the intensive EBM program as well.

Outcomes

The primary outcome was EBM behaviour, measured as guideline adherence and information-seeking behaviour in the intervention and control group during follow-up. Secondary outcomes were EBM attitude and EBM knowledge.

Table 1, Description of the EBM training programs as taught to GP trainees from the GP specialty training in Utrecht, assigned to the intervention or control group.

	Intervention *Integrated EBM training program* *Theoretical basis: Hoes/Grobbee (10) and Offringa 2008 (11)*	**Control** *Stand-alone EBM training program* *Theoretical basis: Offringa 2008 (11)*
Clinical practice (four days a week*)	Tutorial dialogue with supervisor Every day in practice, one hour Integration of EBM in the "EBM tutorial dialogue", once per week with the goal to integrate evidence into the dialogue, combining the clinical expertise of the supervisor with the EBM skills of the trainee.(6)	Tutorial dialogue with supervisor Every day in practice, one hour No integration of EBM
	Critical appraisal of an article together with the supervisor, once a month. All trainees received an e-mail every 3-4 weeks with suggestions for articles, and were encouraged to discuss other clinically relevant articles. These articles were mainly retrieved from "Behind the headlines" (38) and the (last issue of the) national GP journal, Huisarts & Wetenschap.(39)	No critical appraisal together with the supervisor

Institute (one day a week)	- EBM course, 5 days a year, 2 ½ hours
	- Focus on the last two steps of EBM; translation from evidence into clinical practice
	- Patient-related pre- and post-assignments to perform together with supervisor
	- Exchange of last week's experiences in clinical practice
	  - One hour every week
	  - Integration of EBM for 15 minutes, with one trainee preparing a topic, based on a patient consultation, searching for the evidence-base, and chairing the session
	- Coaching session about EBM attitude at the beginning of the year in which trainees were asked to discuss EBM needs and barriers, thus enabling us to tailor the content of the integrated EBM training program to the needs of the trainees.
	- Presentation of a CAT
		- EBM course, 4 days a year, 2 ½ hours
		- Focus on the first three steps of EBM
		- Pre-assignments about imaginary patients
		- Exchange of last week's experiences in clinical practice
		  - One hour every week
		  - No integration of EBM
		-
		-

Online	Possibility for e-learning expanded and improved Online coaching opportunity	Possibility for e-learning -
Extra	Participation in research (PINET) Possibility for extra EBM course	Possibility to participate in research Possibility for extra EBM course
Supervisors	Two days of EBM training at the institute with the same theoretical basis and focus as the program of the trainees, including joint assignments for the trainee and supervisor in clinical practice, and with sessions about how to best integrate EBM in clinical practice, how to engage EBM in tutorial dialogues, and the importance of being a role model Four times a year (1 hour) EBM training with before and after assignments for the trainee and supervisor in clinical practice	- -

* Based on a full-time working contract

Instruments and Measurements

EBM behaviour: guideline adherence and information-seeking behaviour

Guideline adherence was defined as the extent to which clinical management was in accordance with key recommendations in the professional GP clinical practice guidelines. This was assessed with a validated instrument that measures adherence with clinical practice guidelines on the (23) most prevalent conditions in general practice. The instrument assesses compliance on 59 different management decisions (diagnosis N=17, therapy N=20, referral N=22) as described in 27 different clinical practice guidelines covering the 23 most prevalent conditions.(12) Guideline adherence of the participants at T0 and T1 was scored by two independent researchers (YS and LW) for every decision in clinical management (diagnosis, therapy, and referral). In case of disagreements, a third researcher (JvD) made a final decision. At T2 guideline adherence was scored by one researcher (JvD). Possible scores were non-compliant (-1), debatable (0), and full-compliant or motivated deviation (1). Allocated scores were used to assess guideline adherence in five different ways, as shown in table 2.

Information-seeking behaviour was defined as the extent to which trainees used sources of medical information to address clinical queries they encountered in daily practice.(13) A clinical query was defined as every consultation-related question that was a reason to search for an answer.(14) Potential sources for EBM information included clinical practice guidelines, and pre-appraised and primary bibliographic databases (such as Clinical Evidence, Cochrane, and PubMed). Information-seeking behaviour was quantified in five different ways, as shown in table 2.

Trainees were asked to collect data on guideline adherence and information-seeking behaviour in logbooks (see appendix) during all face-to-face patient consultations, either at the GP office or the patient's home, during the measurement periods (excluding telephone consultations); at T0 (baseline) and T1 (at the end of their 3rd year) during eight consecutive days, at T2 (1 year after graduation) during three days (table 2). They were reminded to participate in person (T0 and T1 only), and by e-mail. They were allowed to report more than one complaint per patient contact and more than one query during each consultation. Data were extracted from either paper or digital logbooks (T0), whereas during follow-up (T1 and T2) all registration was digital.

EBM attitude and knowledge

EBM attitude was defined as the mind-set of trainees as to the principles of EBM. It was assessed using the translated version of the McColl questionnaire, consisting of 7 questions, that was validated for the Dutch setting.(15–17) EBM knowledge was defined as the ability to answer questions on EBM, as taught in the EBM training program and was assessed with a newly developed, validated questionnaire (score 0-50), based on the content of the EBM training program.(18) To minimize the impact of repeated measurements (i.e. repeated use of the same questionnaire), no feedback on performance was reported to trainees after each questionnaire.

Table 2, Instruments and measurements used for EBM behaviour, EBM attitude, and EBM knowledge.

		Instrument	Data	Score per item
EBM behaviour	Guideline adherence	Validated instrument that assesses EBM behaviour of GP trainees by measuring rational guideline adherence during patient consultations, thereby allowing motivated deviation in individual patients on 59 different management decisions (diagnosis N=17, therapy N=20, referral N=22) for 23 conditions as described in 27 different Dutch GP clinical practice guidelines(12)	Copied from the electronic medical records (EMR): gender, age, reason for encounter, medical history, results of physical examination, diagnosis, and treatment or referral	-1 Not in accordance with the guideline and no reason mentioned to deviate from it 0 Debatable if in accordance with the guideline due to insufficient or contradicting information 1 In accordance with the guideline or not in accordance but with rational motivation
	Seeking behaviour	Logbook	The presence of a clinical query, the presence of a search, resource used to search, and retrieval of an answer	1 Yes 0 No
EBM attitude		Validated questionnaire with seven questions on attitude (with a visual analogue scale from 0-100) of the translated version of the McColl questionnaire(15-17)		0-100
EBM knowledge		Preliminary version of the validated questionnaire on EBM knowledge (18)		0-1

* Evidence-based resources: Guidelines, PubMed, and pre-appraised bibliographic databases (such as Clinical Evidence, TRIP, Cochrane).

Evaluation of the intervention

Implementation of the (new) integrated EBM training program was assessed by asking trainees to report their compliance to new elements of the EBM training program anonymously, and by asking them for their opinion about the program.

Sample size

The sample size calculation was based on the expected improvement in information-seeking behaviour. For an increase from 15% to 50% (19–21) at a significance level of

	Measurement		
Outcome	**Score**	**Numerator**	**Denominator**
Complete adherence in patient consultations	0-1	All patient consultations with score 1 on all management decisions in that patient consultation	All patient consultations assessed with the instrument
All decisions in patient management		All management decisions with score 1	All management decisions assessed with the instrument
Diagnostic decisions in patient management		All diagnostic management decisions with score 1	All diagnostic management decisions assessed with the instrument
Therapeutic decisions in patient management		All therapeutic management decisions with score 1	All therapeutic management decisions assessed with the instrument
Referral decisions in patient management		All management decisions on referral with score 1	All management decisions on referral assessed with the instrument
Queries per patient	0-1	All clinical queries	All patient contacts
Performed search per clinical query		All performed searches	All queries
Performed search in evidence-based resources per clinical query*		All performed searches in evidence-based resources	All queries
Retrieved answers per performed search		All retrieved answers	All performed searches
Retrieved answers per performed search in evidence-based resources*		All retrieved answers	All performed searches in evidence-based resources.
EBM attitude	0-100	Sum of all scores	All items scored*100
EBM knowledge	0-50	NA	NA

0.05 (2-sided) with a power of 90% and corrected for clustering (two clusters) with an intraclass correlation coefficient of 0.05, a total sample size of 33 trainees was needed (epicalc2000).(22) As results from studies about information-seeking behaviour are generally heterogeneous, we doubled our sample size. Due to the joint educational programs, the total number of groups and thereby participants was substantially higher than needed based on the sample size calculations.

Statistical analysis

Trainee characteristics at baseline were reported as means with standard deviations (SD), medians with interquartile ranges (IQR), or proportions. The outcome guideline

adherence and information-seeking behaviour for both the intervention and control group were reported as proportions. Relative risks (RRs) with 95% confidence intervals (95% CIs) were calculated to compare the two groups. For EBM attitude and knowledge means with 95% CIs were estimated. The effect of the intervention on the outcomes was assessed using mixed models.(23,24) A random intercept was included to account for cluster randomization. An autoregressive residual (i.e. GEE type) covariance matrix was included to correct for the associations between the outcomes at the different time points.(25) In addition, we adjusted for the preferred type of logbook and the national GP knowledge test score at the beginning of the 3rd year. The analyses were based on the intention to treat principle. Missing data were not imputed, as multilevel analysis accounts for loss-to-follow-up. In a sensitivity analysis we excluded trainees from the analyses when they had logged only a few patients at one or more measurements (the lowest decile, *N*=60 at T0, *N*=17 at T1, and *N*=47 at T2). Statistical analyses were performed using Statistical Package for the Social Sciences version 20.0 (SPSS Inc., Chicago, IL, USA), and SAS version 9.2 (SAS Institute Inc., Cary, North Carolina). Comparisons between groups were considered statistically significant at the $p < 0.05$ level.

Ethical considerations

This study design was assessed by the UMC Utrecht Ethics Committee, and regarded as non-eligible for full informed consent. However, we obtained informed consent from the trainees for the use of data from the logbooks, questionnaires, and educational tests during specialty training. Patient data were reported anonymously.

RESULTS

Of the 82 trainees in the 6 groups approached 79 (96%) participated; 39 were randomly allocated to the intervention group, and 40 to the control group (figure 1). Baseline characteristics in the two groups were comparable (table 3), except for the score on the national GP knowledge test (mean of 49 vs. 40), guideline adherence (47% vs. 68%; table 4), and the proportion of clinical queries (25% vs. 17%; table 4). Seventy-six trainees (96%) kept logbooks at T0, 60 (76%) at T1, and 53 (67%) at T2. Of all the logbooks kept by GP trainees, 81% was digital. In total, trainees collected data about 15,894 patient consultations (mean 72, SD 33, range 6-169 per trainee); 7,614 (48%) at T0, 4,973 (31%) at T1, and 3,307 (21%) at T2. At baseline the preferred type of logbook differed significantly between both groups (16% digitally in the intervention group vs. 79% in the control group, $p<0.05$). Questionnaires on EBM attitude and knowledge were completed by 79 (100%), 66 (84%), and 58 trainees (73%) at the three time-points, respectively. More trainees were lost to follow-up in the intervention than the control group (41% vs. 25% for logbooks, and 33% vs. 20% for questionnaires). Baseline characteristics of trainees who did not complete all 3 measurements (*N*=34) did not differ from those who did (*N*=45).

Figure 1. Flowchart.

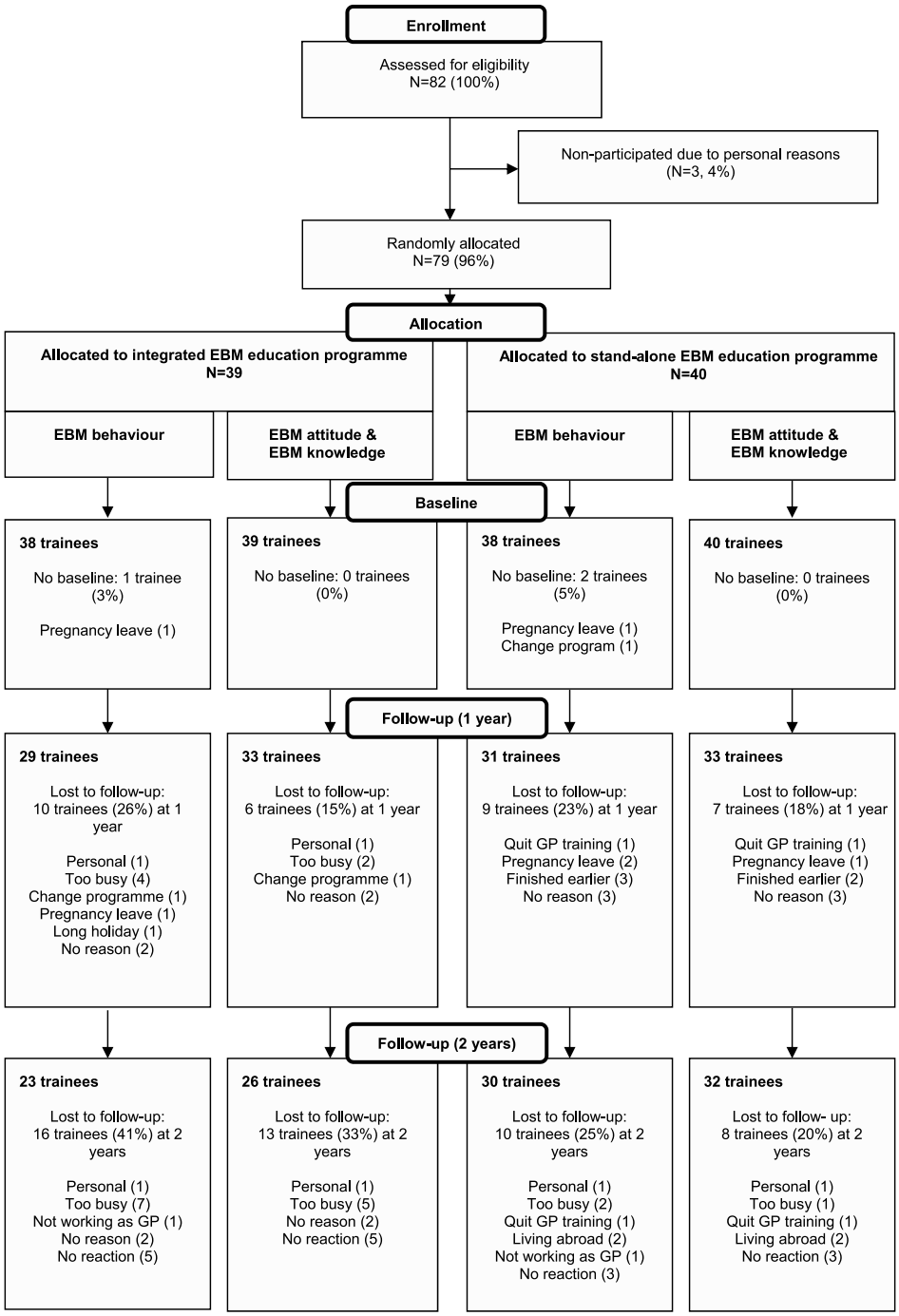


Table 3, Baseline characteristics of third-year GP trainees from the GP specialty training in Utrecht, in 2011 (data collected from questionnaires).

	Intervention N=39 (49%)	Control N=40 (51%)
Demographics		
Age (median, IQR)	31 (29-35)	31 (30-35)
Male (N, %)	14 (36)	7 (18)
Fulltime working (N, %)	29 (75)	25 (63)
Prior medical school		
Same as the GP specialty training (N, %)	20 (51)	21 (53)
Experience		
Time since graduation in years (median, IQR)	4 (3-6)	4 (3-6)
Experience as doctor in months (*median, IQR*)	18 (12-26)	24 (18-37)
Post doctorate qualification (*N, %)*	1 (3)	1 (3)
EBM attitude		
Self-reported, 5-point Likert scale *(median, IQR)*	4 (3-4)	3 (3-4)
EBM knowledge		
Self-reported, 5-point Likert scale *(median, IQR)*	3 (2-3)	3 (2-3)
National GP knowledge test (*mean, SD*)	40 (10)	49 (8)

Intervention effects

Information-seeking behaviour and guideline adherence did not significantly differ between both groups, neither at the end of the third year (T1) nor one year after graduation (T2), with RRs for guideline adherence varying between 0.96 and 0.99 (T1) and 0.99 and 1.10 (T2), depending on the type of management decision, and for information-seeking behaviour varying between 0.97 and 1.16, and 0.90 and 1.10, respectively (table 4). GP trainees motivated why they deviated from guideline instructions in 10% of the patient consultations during training, and in 20% of their consultations after graduating. As for the secondary outcomes mean scores for EBM attitude and knowledge at T1 and T2 did not significantly differ between intervention and control group (table 5). Exclusion of 14 trainees (9 in the intervention and 5 in the control group) who had logged only a few patients at one or more measurements slightly increased the difference between the 2 groups in the proportion of retrieved answers from an RR of 1.05 (0.98-1.13) to a (statistically significant) relative risk of 1.09 (1.02-1.15), but this had no effect on the other outcome measures.

Evaluation of the intervention

Seventeen trainees (44%) in the intervention group reported compliance to the new EBM training elements in daily clinical practice, and evaluated the training program. All respondents reported to have EBM tutorial dialogues with their supervisor, but only 12% did this once a week as suggested, the other less. The majority of responding trainees (53%) did not critically appraise articles with their supervisors, or only less

Table 4. EBM behaviour (95% CI) as recorded in logs by third-year GP trainees from the GP specialty training in 2011.

	T0, baseline^ (N=76)			T1, end of 3rd year^^ (N=60)			T2 one year after graduation^^ (N=53)		
	Intervention N=38	Control N=38	RR	Intervention N=29	Control N=31	RR	Intervention N=23	Control N=30	RR
Information-seeking behaviour									
Queries per patient	0.25 (0.21-0.29)	0.17 (0.15-0.20)	1.45 (1.15-1.83)	0.14 (0.13-0.16)	0.15 (0.13-0.17)	0.95 (0.80-1.12)	0.21 (0.19-0.23)	0.19 (0.16-0.22)	1.10 (0.92-1.32)
Performed search per clinical query	0.77 (0.72-0.82)	0.82 (0.75-0.91)	0.93 (0.83-1.05)	0.71 (0.64-0.79)	0.72 (0.68-0.76)	0.98 (0.87-1.11)	0.60 (0.59-0.61)	0.67 (0.57-0.79)	0.90 (0.76-1.07)
Performed search in evidence-based resources~ per clinical query	0.32 (0.26-0.38)	0.30 (0.24-0.39)	1.05 (0.77-1.44)	0.27 (0.21-0.35)	0.24 (0.15-0.38)	1.16 (0.70-1.91)	0.27 (0.24-0.30)	0.30 (0.23-0.38)	0.90 (0.70-1.16)
Retrieved answers per performed search	0.84 (0.81-0.88)	0.81 (0.77-0.85)	1.04 (0.98-1.11)	0.81 (0.72-0.90)	0.79 (0.72-0.85)	1.03 (0.97-1.10)	0.82 (0.78-0.86)	0.78 (0.71-0.85)	1.05 (0.98-1.13)
Retrieved answers per performed search in evidence-based resources~	0.91 (0.87-0.96)	0.88 (0.84-0.93)	1,03 (0.97-1.10)	0.87 (0.83-0.91)	0.90 (0.82-0.99)	0.97 (0.87-1.07)	0.86 (0.81-0.92)	0.96 (0.89-1.03)	0.90 (0.79-1.03)
Guideline adherence									
Complete adherence on all decisions in a patient consultation	0.47 (0.42-0.53)	0.68 (0.62-0.74)	0.69 (0.60-0.80)	0.62 (0.59-0.66)	0.64 (0.60-0.70)	0.96 (0.86-1.08)	0.67 (0.60-0.74)	0.60 (0.56-0.65)	1.10 (0.97-1.25)
All decisions in patient management	0.69 (0.65-0.73)	0.85 (0.82-0.88)	0.81 (0.76-0.86)	0.83 (0.81-0.84)	0.83 (0.82-0.85)	0.99 (0.97-1.02)	0.83 (0.80-0.87)	0.80 (0.78-0.83)	1.04 (0.98-1.10)
Diagnostic decisions in patient management	0.66 (0.63-0.68)	0.88 (0.86-0.91)	0.74 (0.71-0.78)	0.87 (0.85-0.90)	0.88 (0.87-0.89)	0.99 (0.96-1.02)	0.85 (0.84-0.87)	0.83 (0.79-0.87)	1.03 (0.98-1.09)
Therapeutic decisions in patient management	0.58 (0.51-0.65)	0.71 (0.65-0.78)	0.81 (0.69-0.94)	0.66 (0.62-0.69)	0.66 (0.61-0.72)	0.99 (0.89-1.11)	0.72 (0.68-0.76)	0.66 (0.60-0.73)	1.09 (0.97-1.23)
Referral decisions in patient management	0.81 (0.79-0.84)	0.94 (0.93-0.96)	0.86 (0.83-0.89)	0.93 (0.92-0.95)	0.95 (0.94-0.96)	0.98 (0.96-1.00)	0.91 (0.86-0.97)	0.92 (0.89-0.94)	0.99 (0.92-1.07)

^ Crude data, collected before the intervention at baseline
^^ Adjusted for type of logbook (paper or digital), score on national GP knowledge test, and repeated measures
~ Evidence-based resources: Guidelines, PubMed, and pre-appraised bibliographic databases (such as Clinical Evidence, TRIP, Cochrane).

Table 5, Estimated means (95% CI) of EBM attitude and EBM knowledge of third-year GP trainees from the GP specialty training in 2011.

	T0, baseline^ (N=79)			T1, end of 3rd year^^ (N=66)			T2, one year after graduation^^ (N=58)		
	Inter-vention N=39	Control N=40	Diffe-rence	Inter-vention N=33	Control N=33	Diffe-rence	Inter-vention N=26	Control N=32	Diffe-rence
EBM-attitude									
Validated McColl questionnaire, adjusted for Dutch GP trainees. (15–17) (score 0-100)	46.2 (42.9-49.6)	48.4 (45.6-51.1)	-2.1 (-6.5 - 2.3)	52.5 (49.8-55.2)	51.8 (49.1-54.4)	0.8 (-3.1 – 4.6)	59.5 (56.5-62.5)	58.9 (56.2-61.6)	0.6 (-3.6–4.7)
EBM-knowledge									
Validated EBM knowledge questionnaire.(18) (score 0-50)	29.9 (28.8-31.0)	29.7 (28.5-30.8)	0.2 (-1.4 - 1.9)	29.8 (28.3-31.3)	29.5 (28.0-31.0)	0.3 (-1.8 - 2.4)	30.3 (28.8-31.9)	29.7 (28.2-31.2)	0.6 (-1.6–2.8)

^ Crude data

^^ Adjusted for score on national GP knowledge test, and repeated measures

than once every two months (29%). Nine trainees (53%) conducted one or more e-learning courses during the year, but almost none (12%) had used online coaching. Trainees rated the integrated EBM training program with a mean score of 6 (scale 1-10). About half of the trainees (53%) considered the overall volume of EBM training as adequate. The integration of EBM into clinical practice was rated as sufficient by two third (N=11). Half of the trainees (N=9) reported that their supervisors were not enthusiastic about the new EBM training elements in clinical practice, or that they were not aware of it (N=13). In contrast, more than half (N=10,) felt supported by their supervisors to integrate the EBM training elements into clinical practice. Responding supervisors (N=15) frequently agreed with trainees' opinions.

DISCUSSION

Our results show that an integrated EBM training program in the final 3rd year of the GP specialty training does not improve EBM performance as compared to stand-alone EBM training. EBM behaviour, with guideline adherence (RR 0.96 to 1.10) and information-seeking behaviour (RR 0.90 to 1.16), attitude, and knowledge did not differ between trainees in the two groups, neither during the last year of training, nor in the year following graduation. Overall, guideline adherence among trainees was high in all phases of their training, ranging from 69 to 95%, depending on the type of management decision. In more than half of the consultations trainees adhered to the guideline on all management decisions. GP trainees had on average one clinical query per five patients, and succeeded to retrieve an answer to the vast majority of their questions.

To our knowledge this is the first study performing a randomized controlled trial among GP trainees comparing integrated with stand-alone EBM training, and assessing both short- and long-term effects on EBM behaviour, attitude, and knowledge.(5,26–29) Prior studies assessing the effects of EBM training predominantly used non-controlled designs, and did not include a direct comparison between programs.(5,26,28) Moreover, integration of EBM is poorly defined in the literature, resulting in a large variety in training programs included in studies, while these studies were performed in many different settings and populations.(26,30) The fact that we use clinical examples in our stand-alone EBM training is considered 'integrated' by some already.(30) Prior studies reporting effects of EBM training showed contradictory results. Outcomes most frequently reported were EBM knowledge, skills and attitude.(5,28,31) The few studies assessing effects on EBM behaviour frequently used (retrospective) questionnaires, in contrast to "real-time" measurements as in our study.(29) This may have resulted in an overestimation of the, small, effects of EBM training observed in these earlier studies. The GP trainees' scores on EBM attitude in our study were low compared to earlier studies.(32) The number of clinical queries was lower as well, but the GP trainees very frequently performed a (successful) search.(31) One year after graduation the number of clinical queries in the intervention group was higher than at the end of the third year, but still low in comparison to earlier studies. Although the number of searches in the intervention group decreased after graduation, all GP trainees in our study still performed more searches compared to earlier studies.(31) The effects of EBM training on adherence of individual GP trainees to multiple guidelines for a wide range of management decisions have not been evaluated before. Moreover, we took motivated deviation in individual patients into account, and our results show that trainees appropriately motivated why they deviated from the guideline in 10% of the patient consultations during training, and in 20% of their consultations after graduating.(33–35)

Strengths and limitations

We assessed EBM behaviour in daily clinical practice, using a large number and a wide variety of patient consultations, with a long-term follow up. The choice to perform our study in the third-year of the GP specialty training enabled us to assess the effects of the different EBM training programs until after graduation. However, the high levels of guideline adherence and information-seeking behaviour at the beginning of the third year left little room for improvement, thereby making it very difficult to demonstrate any additional benefit of the intervention. At baseline groups differed in the score on the national GP knowledge test and the preferred type of logbook. Because in our study allocation to intervention or control arm was random, differences between the groups at baseline are random and it is not by definition obligatory to adjust for this.(10) However, as the national GP knowledge measures knowledge of the professional guidelines and therefore may have been correlated to guideline adherence in clinical practice, we accounted for this in our analyses, but this did not materially change our findings.

The differences in preferred type of logbook at baseline had an important impact on the (quality of the) data reported by GP trainees, and thereby on the assessment of guideline adherence with our instrument. For this reason we decided to account for the type of logbook as used at baseline in our analyses, and decided to only collect data digitally at the next two follow-up measurements in both groups. For integration in practice we were dependent on both the GP trainees and the GP supervisors, who are important role models. We know that insufficient skills and knowledge of supervisors, together with a negative attitude, are important barriers for them to use and to teach EBM.(36,37) For this reason we gave them training in EBM as well, and trainees and supervisors performed assignments together.(5) An online survey after the intervention (with a response rate of 44%) showed that only half of these trainees thought that their supervisors were enthusiastic about their new EBM training. Large variability in the way some components of the intervention were delivered in the primary care practices may have diluted a potential effect of our intervention. Although such "lack of adherence" may have reduced the effect of the intervention, such variability is likely to occur in any training program in daily teaching environments. Furthermore, the contrast between both groups may have been limited because our stand-alone EBM training could be considered as "integrated" to some extent, because we do use clinical examples.(30) Comparison with "no EBM training" could have provided more information about the effect(s) of EBM training. However, as EBM is one of the competencies obligatory in the GP specialty training in the Netherlands, this was no option. The time investment needed for GP trainees to collect data in logbooks was substantial. This may have reduced commitment from GP trainees and this may – to save time - have resulted in recording (too) little information on consultations (especially in the paper logbooks where they could not copy from the electronic medical records (EMRs)), underreporting, selective reporting of the number of clinical queries (e.g. only when trying to pursue an answer), or higher rates of loss-to-follow-up. Bias due to the high rates of loss-to-follow-up cannot be fully excluded, although a multilevel analysis like we performed is very robust for (randomly) missing follow up measurements.(23) Moreover, while accounted for in our analyses, the differences between both groups in logbook preference at baseline may have affected the results. Using (copied) data from the EMRs may not have adequately reflected guideline adherence, and misclassification may have occurred due to lack of relevant information. Although guideline adherence was scored by different raters at the three different measurements, we do not expect this to affect the results as the instrument we used has been proven to be reliable, with interrater reliabilities ranging from moderate to excellent. Finally, for practical reasons we randomly allocated the participating trainees at the group level, and we randomly assigned the integrated and stand-alone EBM intervention to the first and the last block of three groups. Although this may theoretically not be the optimal randomization procedure, we do not think that the starting moment of the group was in any way related to clinical capacity of the trainees EBM behaviour.

Conclusion and implications for future

In a randomized controlled trial we found no differences in EBM behaviour, attitude or knowledge between GP trainees who received EBM training through a stand-alone module or those who followed an integrated EBM training. We think that the main explanation is that the stand-alone program adequately increases EBM skills, leaving little room for improvement. EBM skills generally develop well during both trainings: GP trainees in both groups showed consistently high levels of guideline adherence, frequently performed searches in case of clinical queries, that were usually (considered) successful. If EBM training is not yet incorporated in specialty training, we suggest using the integrated EBM training format, as this links to the clinical development of the trainee. Future studies could focus on better integration of EBM training into clinical practice, for instance by engaging supervisors more, as they are important role models for trainees. Moreover, assessment of EBM behaviour at point of care might be made easier, for example by automatically extracting patient consultation data from EMRs.

REFERENCES

1 Sackett DL, Rosenberg WMC, Gray J a M, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. 1996. BMJ. 1996 Feb;312(71).

2 Frank JR. The CanMEDS 2005 physician competency framework. Better standards. Better physicians. Better care. Ottawa R Coll Physicians Surg Canada. 2005.

3 Dawes M, Summerskill W, Glasziou P, Cartabellotta A, Martin J, Hopayian K, et al. Sicily statement on evidence-based practice. BMC Med Educ. 2005 Jan 5;5(1):1.

4 Ilic D, Maloney S. Methods of teaching medical trainees evidence-based medicine: A systematic review. Med Educ. 2014;48(2):124–35.

5 Coomarasamy A, Khan KS. What is the evidence that postgraduate teaching in evidence based medicine changes anything? A systematic review. BMJ. 2004;329(7473):1017.

6 Taylor R, Reeves B, Ewings P, Binns S, Keast J, Mears R. A systematic review of the effectiveness of critical appraisal skills training for clinicians. Med Educ. 2000;34(2):120–5.

7 Van Dijk N, Hooft L, Wieringa-de Waard M. What are the barriers to residents' practicing evidence-based medicine? A systematic review. Acad Med. 2010;85(7):1163–70.

8 Vermeulen MI, Kuyvenhoven M, Zuithoff NPA, van der Graaf Y, Pieters HM. Uitval en stagnatie in de opleiding tot huisarts. Ned Tijdschr Geneeskd. 2011;155:A2780.

9 Van Leeuwen YD, Pollemans MC, Mol SS, Eekhof JA. The Dutch Knowledge Test for General Practice. European Journal of General Practice. 1995. p. 113–7.

10 Grobbee D, Hoes A. Clinical Epidemiology: Principles, Methods and Applications for Clinical Research. 1st ed. London: Jones and Bartlett Publishers. 2009.

11 Offringa, W.J.J. Assendelft RJPMS. Inleiding evidence-based medicine. Klinisch handelen gebaseerd op bewijsmateriaal. M. Bohn Stafleu van Loghum. 2008.

12 Kortekaas, MF, Bartelink, MEL, Van der Heijden, GJMG, Hoes, AW DWN. Development and validation of a new instrument measuring EBM behaviour in clinical practice. submitted.

13 Ely JW, Osheroff J a, Chambliss ML, Ebell MH, Rosenbaum ME. Answering physicians' clinical questions: obstacles and potential solutions. J Am Med Inform Assoc. 2005;12(2):217–24.

14 Kortekaas MF, E M-L, Bartelink L, Boelman L, Hoes AW, De Wit NJ. General practice trainees' information searching strategies for clinical queries encountered in daily practice. Fam Pract. 2015;(pii: cmv046. [Epub ahead of print]):1–5.

15 McColl A, Smith H, White P, Field J. General practitioner's perceptions of the route to

evidence based medicine: a questionnaire survey. BMJ. 1998 Jan 31;316(7128):361–5.

16 Knops AM, Vermeulen H, Legemate DA, Ubbink DT. Attitudes, awareness, and barriers regarding evidence-based surgery among surgeons and surgical nurses. World J Surg. 2009;33(7):1348–55.

17 Te Pas E, van Dijk N, Bartelink MEL, Wieringa-De Waard M. Factors influencing the EBM behaviour of GP trainers: A mixed method study. Medical Teacher. 2012. p. e1–8.

18 Kortekaas MF, Bartelink MEL, de Groot E, Korving H, de Wit NJ, Grobbee DEGD, et al. Development and validation of a new questionnaire to measure EBM knowledge: the Utrecht questionnaire on knowledge on Clinical epidemiology for Evidence-based Practice (U-CEP). submitted.

19 Khan KS, Dwarakanath LS, Pakkal M, Brace V, Awonuga A. Postgraduate journal club as a means of promoting evidence-based obstetrics and gynaecology. J Obstet Gynaecol. 1999;19(3):231–4.

20 Magrabi F, Coiera EW, Westbrook JI, Gosling a S, Vickland V. General practitioners' use of online evidence during consultations. Int J Med Inform. 2005 Jan;74(1):1–12.

21 Van Duppen D, Aertgeerts B, Hannes K, Neirinckx J, Seuntjens L, Goossens F, et al. Online on-the-spot searching increases use of evidence during consultations in family practice. Patient Educ Couns. 2007 Sep;68(1):61–5.

22 Epicalc [Internet]. Available from: http://www.brixtonhealth.com/epicalc.html

23 Hedeker & Gibbons: Longitudinal Data Analysis, Wiley, 2006

24 Groenwold RH1, Moons KG, Peelen LM, Knol MJ, Hoes AW. Reporting of treatment effects from randomized trials: a plea for multivariable risk ratios. Contemp Clin Trials. 2011 May;32(3):399-402. doi: 10.1016/j.cct.2010.12.011. Epub 2010 Dec 31.

25 Zeger SL, Liang KY, Albert PS. Models for longitudinal data: A generalized estimating equation approach. Biometrics. 1988;44(4):1049–60.

26 Young T, Rohwer A, Volmink J, Clarke M. What are the effects of teaching evidence-based health care (EBHC)? Overview of systematic reviews. PLoS ONE. 2014. p. e86706.

27 Oude Rengerink K, Zwolsman SE, Ubbink DT, Mol BWJ, van Dijk N, Vermeulen H. Tools to assess evidence-based practice behaviour among healthcare professionals. Evid Based Med. 2013 Aug;18(4):129–38.

28 Flores-Mateo G, Argimon JM. Evidence based practice in postgraduate healthcare education: a systematic review. BMC Health Serv Res. 2007;7:119.

29 Shaneyfelt T, Baum KD, Bell D, Feldstein D, Houston TK, Kaatz S, et al. Instruments for evaluating education in evidence-based practice: a systematic review. JAMA. 2006 Sep 6;296(9):1116–27.

30 Brauer DG, Ferguson KJ. The integrated curriculum in medical education: AMEE Guide No. 96. Med Teach. 2015;37(4):312–22.

31 Del Fiol G, Workman TE, Gorman PN. Clinical questions raised by clinicians at the point of care: a systematic review. JAMA Intern Med. 2014;174(5):710–8.

32 Te Pas E, van Dijk N, Bartelink MEL, Wieringa-de Waard M. Factors influencing the EBM behaviour of GP trainers: A mixed method study. Med Teach. 2013; 35(3):e990-e997. doi: 10.3109/0142159X.2012.733044. Epub 2012 Oct 26.

33 Pijnenborg L, Braspenning JCC, Kallewaard MBM. Indicatoren op basis van richtlijnen. 2004;305–14.

34 Van den Berg MJ, de Bakker DH, Spreeuwenberg P, Westert GP, Braspenning JCC, van der Zee J, et al. Labour intensity of guidelines may have a greater effect on adherence than GPs' workload. BMC Fam Pract. 2009;10:74.

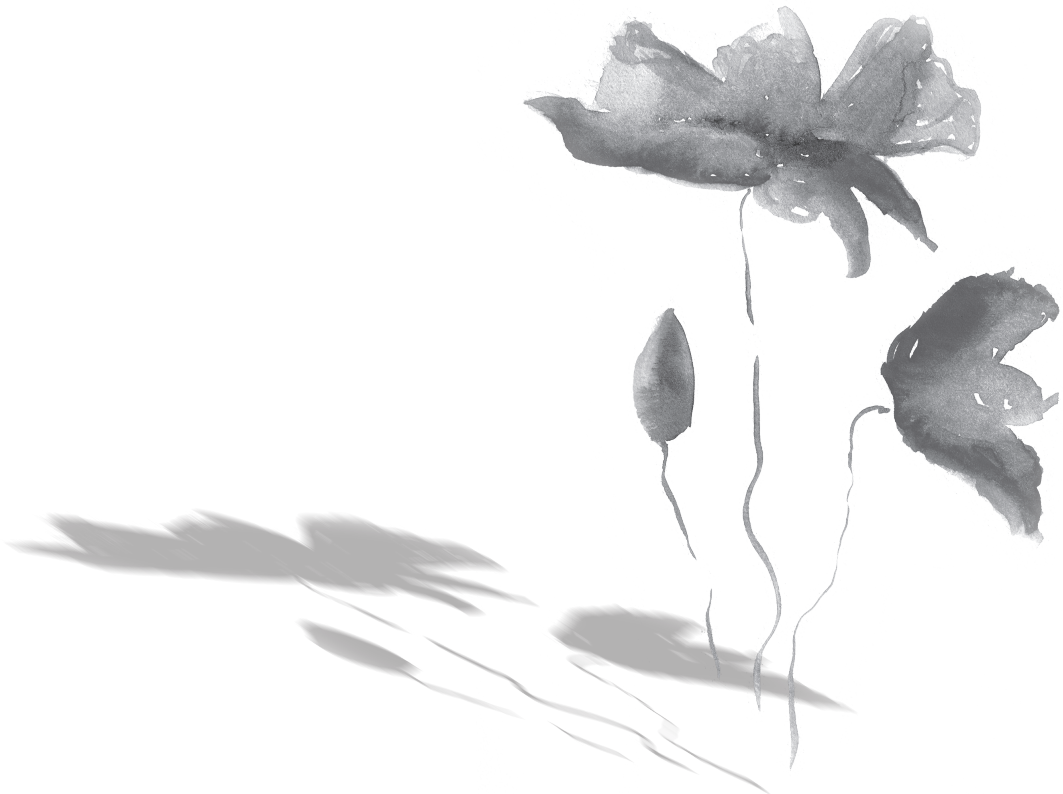
35 Willekens M, Giesen P, Plat E, Mokkink H BJG. Quality of after-hours primary care in The Netherlands: adherence to national guidelines. BMJ Qual Saf. 2011;20(3):223–7.

36 Mayer J, Piter man L. Family Practice-1999-Mayer-627-32. Fam Pr. 1999;16:627–32.

37 Van Dijk N, Hooft L, Wieringa-de Waard M. What are the barriers to residents' practicing evidence-based medicine? A systematic review. Acad Med. 2010 Jul;85(7):1163–70.

38 Behind the Headlines [Internet]. Available from: http://www.nhs.uk/news/Pages/NewsIndex.aspx

39 Huisarts & Wetenschap [Internet]. Available from: http://www.henw.org/



CHAPTER 9

General discussion

"Zou je dat nou wel laten?"
Arnout Visscher

The aims of this thesis were to gain more knowledge on how to improve evidence-based general practice. We first focused on the match between the current evidence base in general practice and the knowledge needs from clinical practice (practice-based evidence). Next, we focused on training in evidence-based medicine, by studying the information needs and seeking behaviour of general practice (GP) trainees, and comparing the effects of two different EBM training programmes, integrated vs. stand-alone, on outcomes relevant to daily clinical practice measured with reliable and valid instruments.

The main findings of this thesis can be summarized as follows:

- The bibliometric analysis we performed on the quantity, themes, and quality of randomized controlled trials (RCTs) in general practice shows that the number of RCTs in the general practice domain published worldwide has tripled in the past 20 years. The number of RCTs from the Netherlands has even increased fivefold. However, the themes of this research correspond insufficiently with the clinical problems presented in daily practice.
- The nationwide database of on-going general practice research projects in The Netherlands we constructed shows that there is a substantial body of on-going research in general practice in the Netherlands, covering a wide range of topics. The majority of this research has a clinical perspective, a high quality research design (RCT), and is non-commercially funded. Also the topics of these projects do not adequately match with the major reasons for encounter in the Dutch general practice; especially clinical questions in dermatology, ophthalmology, reproductive medicine, and gastro-enterology are insufficiently addressed.
- The systematic review we performed on the information needs and seeking behaviour of GP trainees shows that studies are very heterogeneous, with a large variety in design, setting, participants, sample size, instruments, and outcome data. Both GP trainees and trainees in a hospital specialty frequently encounter clinical queries in daily clinical practice, perform searches in the vast majority, and retrieve answers in most cases. They use colleagues as their primary source of information. Trainees mostly perform searches during patient consultations, GP trainees more often than trainees in a hospital specialty. Most queries regard therapeutic questions. The trainees consider most of the answers they find as relevant, i.e. they do change patient's management.
- The information needs and seeking behaviour of third-year GP trainees we observed correspond quite well with results from previous studies. Trainees try to pursue an answer in the vast majority of clinical queries and retrieve an answer for most of the searches performed. However, they seem to have less, i.e. only one in five patients, clinical queries per patient than reported earlier. Next to colleagues, their second most frequently used source of information is the national professional GP guideline database. Other evidence-based Internet resources such as primary or pre-appraised research are used very infrequently.

- The Utrecht questionnaire on knowledge on Clinical epidemiology for Evidence-based Practice (U-CEP) adequately measures knowledge on clinical epidemiology and focuses on aspects relevant to daily clinical practice. It is a valid and reliable evaluation tool for EBM training that can be used for clinicians and trainees.
- The instrument that we developed to assess adherence to professional guidelines proved to be valid, reliable and feasible. It facilitates assessment of three relevant aspects of clinical management decisions (diagnosis, therapy, and referral) for a wide variety of disorders. It can be used to monitor clinical performance and personal EBM learning curve during professional training.
- Overall, adherence to guidelines among GP trainees is high in all phases of their training, ranging from 69 to 95%, depending on the type of management decision. In more than 50% of the consultations trainees adhere to the guideline for all major management decisions.
- In an RCT an integrated EBM training programme for third year GP trainees did not improve EBM performance (as measured by guideline adherence and information seeking behaviour) as compared to stand-alone EBM training.

Improving evidence-based general practice; what is needed?

The term evidence–based medicine (EBM), originating from the work of Pierre Louis in Paris in the 19th century, was introduced as the cornerstone of today's medical practice only 20 years ago. EBM is defined as the integration of clinical expertise, patient values and the best available clinical evidence in daily clinical practice.(1) It requires knowledge, skills, and a positive attitude towards EBM to be well applied in clinical practice. Since the introduction of EBM, its impact on clinical practice has evolved greatly. Still, it has not been completely incorporated in daily clinical practice of GP trainees, as our results show. Practising EBM demands GP trainees to be critical and 'ask' for evidence, and requires researchers to supply them with the evidence needed. Hence, to improve the practice of EBM in daily clinical practice we both have to ensure an adequate portfolio of practice-based research (i.e. providing the evidence needed) and train future general practitioners (GPs) to apply the resulting evidence in evidence-based medicine. In the next paragraphs we discuss ways to improve evidence-based general practice.

More focus on clinically relevant research questions

Although there is a substantial body of research in the field of general practice covering a wide range of topics, the research themes correspond insufficiently with the clinical problems encountered in daily practice (***chapter 2 and 3***). To improve the practice of EBM it is important that research projects better address clinically relevant questions. This so-called practice-based research in general practice is essential as it can fill knowledge gaps and provides the fundamentals to integrate evidence into daily clinical practice, combining it with clinical expertise and patient values, thereby practising EBM.(2)

More insight in knowledge gaps could help researchers to better focus on relevant clinical issues in day-to-day practice, and cover topics that are not adequately addressed by governmental and private funding organisations. Presently they primarily focus on chronic diseases, frequently addressing diseases such as diabetes mellitus, hypertension, depression, and asthma. They focus much less on clinical problems of, for example, the skin, eyes, and the digestive tract, despite their high incidence in clinical practice (***chapter 2***). Furthermore, multimorbidity receives insufficient attention in primary care research.

In 2001 the Dutch College developed an overview of knowledge gaps in the national GP guidelines. Unfortunately, 10 years later most researchers were still not aware of this overview (***chapter 3***). In 2014 the College developed a web-based tool for this overview, in connection with the GP guidelines, to improve the use of this system. It now also provides an overview of on-going research projects in the field of general practice in the Netherlands, based on our overview (https://www.nhg.org/lacunes). Researchers are requested to add new research projects. If adequately used, this tool can be a major step forward in linking future research themes to knowledge gaps in general practice. A limitation of this system is that knowledge gaps outside the scope of the 96 currently available guidelines are not included in this database.(3)

Governmental funding organizations, such as the Dutch ZonMw (the Netherlands Organisation for Health Research and Development), are key actors in monitoring and driving the research fields, and in linking the needs of patients, health care professionals, and governmental bodies. Until recently the ZonMw programme 'Common disorders' (Dutch: 'Alledaagse ziekten') provided Dutch researchers with a unique funding opportunity to study symptoms and diseases commonly seen in general practice.(4) Despite its relatively modest budget, the funding programme made an important contribution in expanding the research field of common disorders in primary care. Thirty-four projects led to 89 publications of research findings in prestigious international journals (mean 3.6 per completed project), 29 in Dutch journals, and the results of 9 projects were rapidly incorporated in professional guidelines.(4) This reflects both the scientific importance and practical relevance of the funded projects, and therefore sets a good example of how to stimulate practice-based research. Given the inability of the main stream funding programmes to do so, it is important to continue such unique funding opportunities. A dedicated research programme of SBOH-ZonMW dedicated to research in general practice, medical education, and elderly medicine aims to fill this gap. The programme started in 2013 and is funded by the SBOH, i.e. the government funded organisation that facilitates and finances the GP specialty and elderly medicine training in the Netherlands.(5) Despite the positive evaluation of its first three years, continuation of the programme is not yet certain.

Improving the fit between research methodology and clinical practice

Apart from addressing questions that meet clinical practice needs, research should also apply designs, include patients, and study determinants and outcomes that match clinical

practice. Research should be of high quality, but pragmatic as well, to account for the variability in daily clinical practice and thereby enhancing the generalizability of the results.(6) Although RCTs are considered the gold standard to quantify causal associations, they do not always allow optimal evaluation of the effects of complex interventions, such as chronic care, prevention, or multidisciplinary intervention programmes.(7) Heterogeneity of patients, the need for personification of the intervention, and the required positive attitude of professionals often hampers a standardised implementation of the intervention. This sometimes challenges researchers to use more pragmatic designs, such as quasi experiments with pre-post comparison, or stepped wedge designs, albeit that these designs do have methodological limitations.(7) Participants should ideally resemble the clinical practice patient population, not applying strict inclusion criteria, especially when the characteristics on which study participants and the typical practice population differ modify the effect of the determinants on the outcome. Interventions under study, diagnostic as well as therapeutic, should be feasible and fit in clinical practice, and outcomes should be pragmatic and relevant to both patients and health care professionals. Studies should not only focus on statistically significant results, but – more importantly - on results that are considered by many to be clinically relevant. To improve interpretation of research findings, both quantitative and qualitative outcomes should be addressed, not only focussing on the (possible) effects, but also on the perception of patients and professionals, and on the (possible) harms caused. Finally, research findings should be reported in a way that clearly demonstrates their relevance to clinicians and patients. For example, for some health care professionals the number needed to treat (i.e. 1/risk difference) or number needed to screen in diagnostic research are intuitively attractive to work with, although many patients find these numbers hard to understand.(8)

Ensure that guidelines are easily applicable to clinical practice

The national GP guidelines are the main source of evidence-based information for GP trainees in clinical practice (***chapter 5***). With, around, hundred highly appreciated guidelines the Dutch College has created an excellent knowledge base for the practice of EBM. However, the application has limitations. First, they do not cover all clinical problems. Moreover, they are based on the 'average' patient and therefore do not fit every patient. Currently, guidelines do not ensure that individual values and preferences are accounted for in management decisions.(9) This makes it difficult to personalize recommendations, even for those skilled in EBM.(10) Personalised guideline interpretation may solve this.(9) Furthermore, the need to individually assess the applicability of a recommendation in a guideline in a specific patient, should be emphasized stronger, both in the complete guidelines, as in the short summaries.(11) The quality of evidence underlying recommendations should be reported in conjunction with the conclusions, as is done in most Dutch multidisciplinary (i.e. GPs and other medical specialties) evidence-based guidelines.(12) Additionally, in the light of multimorbidity, reflections on possible interferences with recommendations in other guidelines could improve the applicability of guidelines. Currently, most guidelines are

written as though patients only have one condition, whereas multimorbidity is the norm, especially in older patients.(13) Furthermore, it should be reported where uncertainty makes generalizability of recommendations difficult.(10) Guidelines should explicitly mention uncertainties, such as those related to comorbidity.(10) Patient decision support tools could help to individualize decisions for patients and to facilitate shared decision making (SDM), by comparing treatments and explicitly explain the possible benefits and harms of an intervention.(9) In this light, the website thuisarts.nl (https://www.thuisarts.nl/), that allows patients to easily and quickly find medical information on diseases and health, could be improved by adding such information.(14)

Teach health care professionals that EBM is the optimal balance between evidence, expert opinion and individual patient needs

At present there are still major misinterpretations about EBM.(11) EBM is frequently considered to consist of critical interpretation of evidence only. For years, textbooks and courses on EBM predominantly addressed the aspect 'evidence', whereas the aspects 'clinical expertise' and 'patient values' were given relatively little attention.(15) It is important to emphasize that EBM is really about the integration of evidence with clinical expertise and patient values. We have to change the frequently encountered attitude of 'clever nihilism', because of too much focus on the methodological problems encountered in published studies, leading to scepticism towards EBM. Instead we should adopt an attitude of 'mature pragmatism', i.e. reasonably balancing EBM critical skills with the exigencies of clinical practice.(16) It should be stressed that EBM is not a solution for undifferentiated clinical challenges that are difficult to manage in day-to-day practice, but merely a way to help GP trainees making knowledgeable, well-founded medical decisions in clinical practice.(17)

Improve training in EBM

The results of our RCT, comparing an integrated EBM training programme in the final third year of the GP specialty training with a stand-alone EBM training programme, show that integrating EBM training had no beneficial effect compared to stand-alone EBM training (***chapter 8***). One of the possible explanations for these findings is that both programmes we compared in our clinical trial are effective. Indeed, GP trainees tried to pursue an answer in the vast majority of clinical queries (***chapter 5***), retrieved an answer for most of the searches performed (***chapter 5***), and showed high levels of guideline adherence, while taking motivated deviation into account (***chapter 8***). Despite these positive EBM findings, there still is ample room to improve EBM training.

Ask

GP trainees in our study did not have as many clinical queries per patient as expected (***chapter 5***). Apparently, the ability to identify a clinical problem can be improved. Ideally, we want to teach GP trainees to always be critical towards their medical performance. For this it is important they (learn to) acknowledge that knowledge

constantly evolves, and that knowledge acquired during medical school or the GP specialty training requires regular updating. It should be one of the top priorities of an EBM training programme to make both GP trainees and their supervisors aware of this.

Acquire

GP trainees should be more efficient in choosing their knowledge database in case of questions. Collegial advice is their primary source of information, providing answers more quickly than other sources of information(18), but often being more experience- than evidence-based.(19) It would be better if GP trainees search in evidence-based sources of information, such as the national GP guidelines, more. To facilitate this we developed a search strategy, the so-called 'EBM highway', which very much aligns with the preferred evidence-based sources of information used by GP trainees (***chapter 5***), and includes an easily applicable hierarchy, based on the quality of the information included. Furthermore, we should teach GP trainees how to rate information retrieved from other internet-based sources of information, such as Google. It may be that this type of information seeking behaviour of GP trainees is more efficient in daily clinical practice, and therefore we should teach how to find valuable information using this approach.(20) As stated in the Future Vision 2022 document of the College of General Practice in the Netherlands, optimisation of the availability of evidence-based resources during patient consultations by further integrating clinical decision support systems in the electronic medical records could help to better match information needs and save time as well.(21)

Appraise

In EBM training trainees frequently learn appraisal skills by critically evaluating primary evidence, such as an original article. I think it would meet trainees' needs better to teach these skills using the evidence-based sources of information most frequently used in clinical practice; i.e. the guidelines. Skills could be taught by interpreting the information provided on the quality of evidence and reasoning behind recommendations in guidelines. Skills should be taught at the level of an EBM user instead of an EBM doer, focusing on clinical practice, and on issues relevant to clinical practice.(22,23) Moreover, discussions on the reasoning behind recommendations and the quality of underlying evidence could be integrated in (regular) group sessions at the trainees' institute where which clinical issues are taught and discussed.

Apply

One of the main differences between the stand-alone and integrated EBM training programme we compared in our randomized controlled trial was the focus on the application of EBM. This step is very important, as it helps trainees to combine the evidence as retrieved with their clinical expertise, translate it to the patient, and combine it with the patient values. To improve the applicability of EBM, the principles of EBM could be combined with the principles of SDM. Clinicians and patients could then share the best available evidence (presented in a patient-friendly summary)

to make joint decisions, and clinicians could support patients to consider options, to achieve informed preferences.(24) EBM and SDM could be complementary. Integration of EBM and SDM training would stress the importance of the last step of EBM, integrating the evidence as retrieved with patient values.

Assess

For monitoring of EBM performance assessment of the degree to which GP trainees work in accordance with guidelines could be used to start discussions with fellow GP trainees on when, why and how to over-ride recommendations in the guidelines. This could enhance the awareness among GP trainees that EBM really is about the *integration* of evidence with clinical expertise and patient characteristics and values. The instrument we developed could facilitate this, as it can be used to assess guideline adherence, incorporating motivated deviation (***chapter 7***).

The role of GP supervisors

Present GP supervisors are not often the best role model for practising EBM.(25) They are not optimally trained in practising EBM, and are less accustomed to using internet based resources than trainees.(26) However, it is important to emphasize that GP supervisors are role models in *all* competencies required at the end of the GP specialty training, EBM being one of them. They could play an important role in encouraging trainees to apply EBM. While they may not always feel competent in acquiring and appraising evidence (18), they do have (plenty of) clinical expertise and know their patients well. The supervisor's clinical expertise with trainee's knowledge and skills on clinical epidemiology could reinforce each other. Based on this hypothesis we asked trainees and supervisors in the integrated EBM training programme of our RCT to integrate EBM into the "tutorial dialogue" once a week. The implementation varied considerably. In hindsight, we probably underestimated how difficult it was for both trainees and supervisors to integrate EBM training into tutorial dialogues in clinical practice, when one (or both) feel(s) incompetent. In future, we should support them better.

Trainees have the future, also in evidence-based general practice

To improve evidence-based general practice it is important to continue redeveloping the training of EBM. For the future of general practice it is important that future GPs have a critical attitude, are well equipped to stay up to date, and make evidence-based decisions on how to optimally manage their individual patients. We have to encourage GP trainees to be curious and eager for new information, and equip them with the tools needed to find, appraise and apply this information.(27) Improving evidence-based general practice requires all key actors to participate. Together we can take away the misunderstanding that EBM is about evidence only, take away the attitude of 'clever nihilism', and focus on the integration of evidence with clinical expertise and patient values in a mature, pragmatic way.

A clinical query about a patient with a Bell's palsy... 25 years from now

Twenty-five years from now, at the beginning of the 2040s, a patient enters the waiting room for an appointment with his general practitioner (GP). Since that morning he suddenly can't close his mouth and lost control of the muscles on the left side of his face. Since he received an e-message that the appointment with the GP was delayed 15 minutes, he arrives later and does not have to wait: a middle-aged GP calls him to her office. The GP asks about his complaints, and performs a physical examination. After a while the GP says she thinks it is a Bell's palsy, and has to check which treatment option is most optimal for this patient.

Within the EMR a search engine searches for information on this topic in the guideline, compares it to professional guidelines from other specialties, such as ENT surgeons, and searches recent literature on the subject. These searches result in the conclusion that in case of a Bell's palsy it is best to give both antiviral medication and corticosteroids to reduce the chance of recurrent Bell's palsy, but that this therapy has no effect on the recovery process.

The GP and patient discuss this and decide to prescribe both the antiviral medication and corticosteroids. The patient receives all information on his electronic patient record in his personal "health cloud", where evidence underlying this decision is presented in a patient-friendly manner.

REFERENCES

1 Sackett DL, Rosenberg WMC, Gray J a M, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. 1996. BMJ. 1996 Feb;312(71).

2 de Wit N. Huisartsgeneeskunde: van alledaagse romantiek naar academische praktijk. Oratie Niek de Wit. Universiteit van Utrecht; 2010.

3 nhg.org [Internet]. Available from: https://www.nhg.org

4 Van Der Wouden JC, Kortekaas M, Dreesens D, Keuken D, Dekker J. Alledaagse ziekten: wat heeft 10 jaar subsidie opgeleverd? Huisarts Wet. 2013;56(5):202–7.

5 SBOH Z. Huisartsgeneeskunde en Ouderengeneeskunde onderzoek in de klinische huisartsgeneeskunde, de klinische ouderengeneeskunde en de opleidingen voor het specialisme huisarts en specialist ouderengeneeskunde. Addendum bij Programma Kwaliteit van Zorg: Versnelle [Internet]. Den Haag; 2013. Available from: http://www.zonmw.nl/uploads/tx_vipublicaties/Programmatekst_HGOG.pdf

6 Thabane L, Kaczorowski J, Dolovich L, Chambers LW, Mbuagbaw L. Reducing the confusion and controversies around pragmatic trials: using the Cardiovascular Health Awareness Program (CHAP) trial as an illustrative example. Trials [Internet]. 2015 Jan;16:387. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4557925&tool=pmcentrez&rendertype=abstract

7 Grobbee D, Hoes A. Clinical Epidemiology: Principles, Methods and Applications for Clinical Research. 1st ed. London: Jones and Bartlett Publishers. ; 2009.

8 Gigerenzer G, Gaissmaier W, Kurz-Milcke E, Schwartz LM, Woloshin S. Helping Doctors and Patients Make Sense of Health Statistics. Psychol Sci Public Interest [Internet]. 2007 Nov;8(2):53–96. Available from: http://www.ncbi.nlm.nih.gov/pubmed/26161749

9 Elwyn G, Quinlan C, Mulley A, Agoritsas T, Vandvik PO, Guyatt G. Trustworthy guidelines - excellent; customized care tools - even better. BMC Med [Internet]. 2015 Jan;13:199. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4556022&tool=pmcentrez&rendertype=abstract

10 McCartney M, Treadwell J, Maskrey N, Lehman R. Making evidence based medicine work for individual patients. BMJ [Internet]. 2016 Jan;353:i2452. Available from: http://www.ncbi.nlm.nih.gov/pubmed/27185764

11 Greenhalgh T, Howick J, Maskrey N. Evidence based medicine: a movement in crisis? Bmj [Internet]. 2014;348(jun13 4):g3725–g3725. Available from: http://www.bmj.com/cgi/doi/10.1136/bmj.g3725

12 Faber MJ, Burgers JS, Westert GP. A sustainable primary care system: lessons from the Netherlands. J Ambul Care Manage [Internet]. Jan [cited 2016 May 26];35(3):174–81. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22668606

13 Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet [Internet]. 2012 Jul 7 [cited 2014 Nov 5];380(9836):37–43. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22579043

14 thuisarts [Internet]. Available from: https://www.thuisarts.nl/

15 Kelly MP, Heath I, Howick J, Greenhalgh T. The importance of values in evidence-based medicine. BMC Med Ethics [Internet]. 2015 Oct 12 [cited 2015 Oct 13];16(1):69. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4603687&tool=pmcentrez&rendertype=abstract

16 Meserve C, Kalet A, Zabar S, Hanley K SM. Clever Nihilism: Cynicism in evidence based medicine learners. Med Educ Online [Internet]. 2005;10(4). Available from: http://www.med-ed-online.org

17 O'Riordan M, Dahinden A, Aktürk Z, Ortiz JMB, Dağdeviren N, Elwyn G, et al. Dealing with uncertainty in general practice: an essential skill for the general practitioner. Qual Prim Care [Internet]. 2011 Jan [cited 2016 Apr 10];19(3):175–81. Available from: http://www.ncbi.nlm.nih.gov/pubmed/21781433

18 te Pas E, van Dijk N, Bartelink MEL, Wieringa-De Waard M. Factors influencing the EBM behaviour of GP trainers: A mixed method study. Medical Teacher. 2012. p. e1–8.

19 Korownyk C, Ross S, Ma V, Aaron S, Allan GM. A descriptive analysis of a novel intervention

to help residents become evidence users. Med Teach. 2013 Oct;35(10):e1546–50.

20 McCord G, Smucker WD, Selius BA, Hannan S, Davidson E, Schrop SL, et al. Answering questions at the point of care: do residents practice EBM or manage information sources? Acad Med. 2007 Mar;82(3):298–303.

21 Toekomstvisie Huisartsenzorg 2022. Utrecht; 2012.

22 Abdelhamid A. Towards an easier way to practice evidence-based medicine: The doer, user, and replicator modes of practice. Middle East Fertil Soc J. 2010;15(2):119–22.

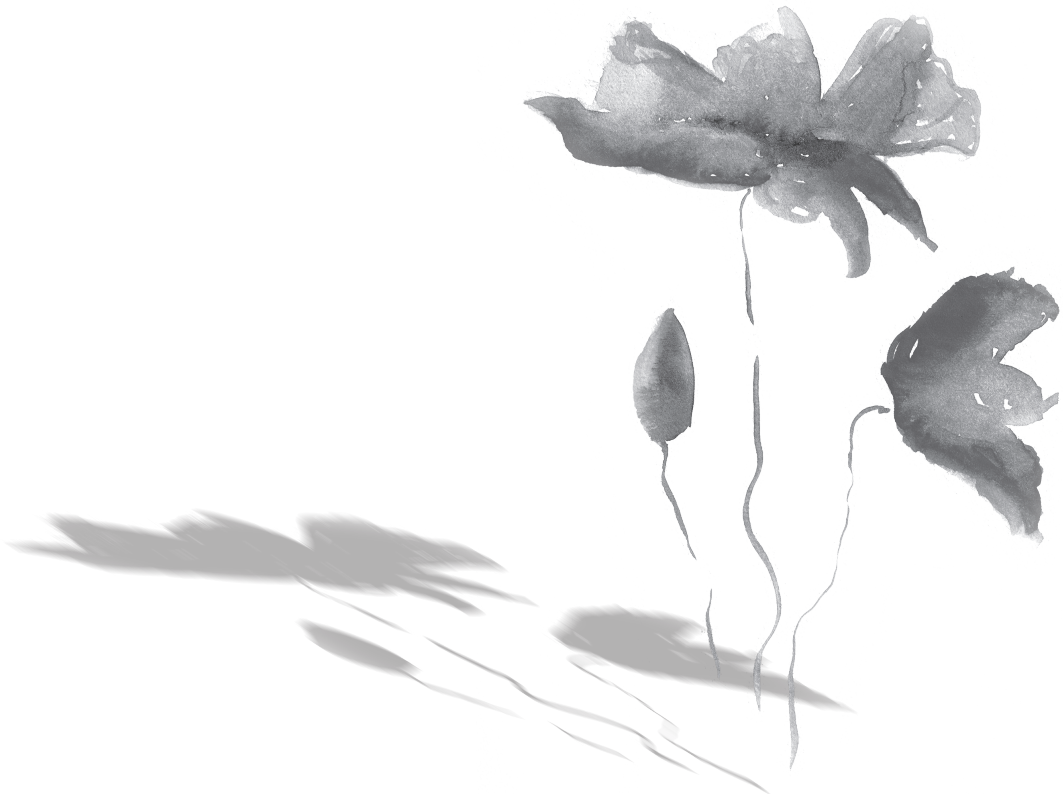
23 Straus SE, Green ML, Bell DS, Badgett R, Davis D, Gerrity M, et al. Evaluating the teaching of evidence based medicine: conceptual framework. BMJ [Internet]. 2004 Oct 30 [cited 2016 May 27];329(7473):1029–32. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=524561&tool=pmcentrez&rendertype=abstract

24 Elwyn G, Laitner S, Coulter A, Walker E, Watson P, Thomson R. Implementing shared decision making in the NHS. BMJ [Internet]. 2010 Jan 14 [cited 2015 Sep 1];341(oct14_2):c5146. Available from: http://www.bmj.com.proxy.library.uu.nl/content/341/bmj.c5146.long

25 van Dijk N, Hooft L, Wieringa-de Waard M. What are the barriers to residents' practicing evidence-based medicine? A systematic review. Acad Med. 2010;85(7):1163–70.

26 Davies K, Harrison J. The information-seeking behaviour of doctors: a review of the evidence. Health Info Libr J. 2007 Jun;24(2):78–94.

27 Horsley T, Hyde C, Santesso N, Parkes J, Milne R, Stewart R. Cochrane Database of Systematic Reviews [Internet]. The Cochrane database of systematic reviews. Chichester, UK: John Wiley & Sons, Ltd; 1996 [cited 2016 Apr 26]. CD001270 p. Available from: http://www.ncbi.nlm.nih.gov/pubmed/22071800



CHAPTER 10

Summary
Samenvatting
Dankwoord
Curriculum Vitae

"Liever spijt van iets dat je wel hebt gedaan,
dan van iets dat je niet hebt gedaan."
Onbekend

SUMMARY

This thesis provides knowledge on how to improve the practice of evidence-based medicine (EBM) by general practitioners (GPs) and GP trainees. EBM is defined as the integration of clinical expertise, patient values and the best available clinical evidence in daily clinical practice.(1) We describe the current clinical evidence base in general practice (GP), both nationally and internationally. Moreover, we describe the information needs and seeking behaviour of trainees. Finally, we compare the effects of two EBM training programmes for GP trainees on outcomes relevant to daily clinical practice; stand-alone versus integrated EBM training. In the stand-alone EBM training programme EBM is trained in theoretical, stand-alone educational sessions. This differs from integrated EBM training, in which EBM is trained in a clinical context (as well), and teaching sessions are based on recent patient consultations in the trainees practice. We measured the effects with feasible, reliable and valid instruments, of which two were developed and validated by us.

Chapter 2 describes a bibliometric analysis on the quantity, themes, and quality of randomized controlled trials (RCTs) in GP. We assessed whether the themes of these RCTs matched with reasons for encounter in the Dutch general practice. It shows that the number of RCTs in the GP domain published worldwide has tripled in the past 20 years, from 48 RCTs in 1990 to 161 in 2009 (those from the Netherlands have even increased fivefold, from 5 to 25, respectively). However, the themes of the research correspond inadequately with the presentation of clinical problems in daily practice. Most RCTs addressed psychiatry (code P; 28%), circulatory disorders (code K, 13%), and general problems (code A, 13%). The United States (28%), the United Kingdom (26%) and The Netherlands (10%) conducted the majority of RCTs.

In ***chapter 3*** we provide a nationwide overview of on-going GP research in the Netherlands. The nationwide database demonstrates that there is a substantial body of on-going research in general practice in the Netherlands, covering a wide range of topics. In 2015 196 research projects were conducted. The majority of this research had a clinical perspective, a high quality research design (RCT, 35%), and was non-commercially funded. However, frequency wise the research topics did not adequately match with the reasons for encounter in the Dutch general practice. Thematically most research projects addressed chronic diseases; mainly cardiovascular risk management (8%), depressive disorders (8%), and diabetes mellitus (7%). Clinical questions in dermatology, ophthalmology, reproductive organs, and the digestive tract were insufficiently addressed.

The systematic review in ***chapter 4*** shows the information needs and seeking behaviour of trainees. The 35 selected studies regarding this topic are very heterogeneously with a large variety in design, setting, participants, sample size, instruments, and outcome data. The results in the review show that both GP trainees and trainees in a hospital-based specialty on average have one clinical query per patient in daily clinical practice,

perform searches in the vast majority (83%), and retrieve answers in most cases (79%). They use colleagues as their primary source of information (25-66%). Trainees primarily perform searches during patient consultations, GP trainees more often than trainees in a hospital-based specialty, but most frequently not in the presence of the patient. Most queries regard therapeutic questions. Answers are considered to change patient's management frequently.

In ***chapter 5*** we report the results of a descriptive study among 76 GP trainees on how and how often they search for answers to clinical queries they encounter in daily clinical practice. We show that the information needs and seeking behaviour of third-year GP trainees correspond quite well to results from previous studies. They try to pursue an answer in the vast majority of clinical queries (80%), and retrieve an answer for most of the searches performed (80%). However, they seem to have less clinical queries per patient as reported earlier, only in one of five patients. Searching takes – on average – four minutes. Although we did not assess whether the trainees' clinical decisions improved patient care, trainees reported that the retrieved information resulted in improved clinical decision making 26% of clinical queries. Next to colleagues (28%) frequently used primary sources of information are the national GP guidelines (26%). Other evidence-based resources (found on the internet), such as primary or pre-appraised research resources, are used very infrequently.

In ***chapter 6*** the development and validation of the Utrecht questionnaire on knowledge on Clinical epidemiology for Evidence-based Practice (U-CEP) is described. The U-CEP measures knowledge on clinical epidemiology, which is a prerequisite for practising EBM in daily clinical practice. The U-CEP consists of two sets of 25 questions each. We show that it is a valid and reliable evaluation tool among different clinicians. It adequately measures knowledge on clinical epidemiology, focusing on aspects relevant to daily clinical practice, and.

In ***chapter 7*** we describe the development and validation of an instrument we developed to assess guideline adherence of GP trainees in clinical practice, while allowing motivated deviation in individual patients. The instrument facilitates assessment of three relevant aspects of clinical management decisions (diagnosis, therapy, and referral) for a wide variety of disorders in daily clinical practice. Evaluation of the instrument among 12,106 patient consultations of 76 GP trainees on 12,587 different reasons for encounter proved the instrument to be valid, reliable and feasible. The instrument can be used to monitor clinical performance and the application of EBM during professional training.

In ***chapter 8*** we describe the results of a cluster RCT we performed among 79 third-year GP trainees, assessing the effects of an integrated EBM training programme compared to stand-alone EBM training. In the stand-alone EBM training programme 40 GP trainees trained EBM in theoretical, stand-alone educational sessions at the

institute. In the integrated EBM training 39 GP trainees were trained in EBM in daily clinical practice as well. Our RCT shows that an integrated EBM training programme in the GP specialty training did not lead to better EBM knowledge, attitude and behaviour among GP trainees, not at the end of their training nor during daily clinical practice one year after graduation. GP trainees in both groups show high adherence (and motivated deviation in individual patients) to guidelines in all phases of their training, ranging from 69% to 95%, and perform a high number of successful (78-84%) searches (60-82%) in case of clinical queries (14-25%).

To conclude, an increasing, substantial body of qualitatively good research in general practice covering a wide range of topics. The themes of this research correspond insufficiently with the clinical problems presented in daily practice. GP trainees frequently search for answers to clinical queries, primarily asking colleagues or using the guidelines. They infrequently search for evidence using evidence-based primary or pre-appraised research resources (found on the internet). Better insight in how GP trainees apply EBM in daily clinical practice can help to improve training in EBM (even further), matching daily clinical practice better.

References

1 Sackett DL, Rosenberg WMC, Gray J a M, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. 1996. BMJ. 1996 Feb;312(71).

SAMENVATTING

In dit proefschrift beschrijven we mogelijke manieren om de toepassing van evidence-based medicine (EBM) door huisartsen en artsen in opleiding (aios) tot huisarts te verbeteren. Onder EBM verstaan we het expliciet, oordeelkundig en consciëntieus gebruikmaken van het beste beschikbare bewijs bij het maken van een keuze voor de beste behandeling van een patiënt.(1) We beschrijven de huidige stand van zaken met betrekking tot de evidence die aanwezig is binnen de huisartsgeneeskunde, nationaal en internationaal. Daarnaast beschrijven we de informatiebehoeften van aios huisartsgeneeskunde en hoe zij hier in de dagelijkse praktijk mee omgaan. We vergelijken twee manieren waarop EBM onderwezen kan worden met elkaar; de reguliere 'stand-alone' manier, en een nieuw ontwikkelde 'geïntegreerde' manier. Bij de reguliere manier wordt het EBM-onderwijs aangeboden op de terugkomdagen, in sessies die los staan van de klinische praktijk. Bij geïntegreerd EBM-onderwijs wordt het onderwijs zoveel mogelijk in de dagelijkse klinische praktijk geïntegreerd. We vergelijken de effecten van beide onderwijsmethoden op uitkomsten die relevant zijn voor de praktijk. Hierbij gebruiken we bruikbare, betrouwbare en valide instrumenten, die deels nieuw door ons zijn ontwikkeld.

In ***hoofdstuk 2*** beschrijven we de resultaten van een bibliometrische analyse waarin we de publicaties van gerandomiseerde gecontroleerde trials (RCT's) in de huisartsgeneeskunde hebben geanalyseerd. We hebben hierbij zowel de thematische ontwikkeling, als de kwantiteit en kwaliteit van de onderzoeken beoordeeld. Ook onderzochten we of de onderwerpen van de RCT's aansloten op de klinische praktijk. De resultaten laten zien dat het aantal gepubliceerde RCT's in de huisartsgeneeskunde de afgelopen twintig jaar flink is gestegen met 48 RCT's in 1990 tot 161 in 2009 (in Nederland van 5 naar 25, respectievelijk), maar dat de verdeling van onderzochte thema's nog niet goed aansluit op de verdeling van de prevalentie van aandoeningen in de dagelijkse huisartspraktijk. De thema's die het meest onderzocht werden waren psychiatrie (28%), hart- en vaatziekten (13%) en algemene zaken (13%). De meeste publicaties waren afkomstig uit de Verenigde Staten (28%), Groot-Brittannië (26%) en Nederland (10%).

In ***hoofdstuk 3*** beschrijven we de resultaten van een inventarisatie van lopend onderzoek binnen de huisartsgeneeskunde in Nederland. De ontwikkelde nationale database laat zien dat binnen de Nederlandse huisartsgeneeskunde veel onderzoek wordt gedaan, anno 2015 lopen er 296 onderzoeken. Het merendeel van deze onderzoeken betreft klinische onderwerpen, is van hoge kwaliteit (RCT, 35%) en wordt vooral uit de tweede en derde geldstroom gefinancierd. De inventarisatie laat zien dat ook in het lopend onderzoek de verdeling van onderzoeksthema's nog niet optimaal aansluit bij de prevalenties van aandoeningen zoals deze in de dagelijkse huisartspraktijk in Nederland gepresenteerd worden. Met name chronische ziekten, zoals hart- en vaatziekten (8%), depressieve klachten (8%) en diabetes mellitus (7%)

worden vaak onderzocht. Dermatologie, oogheelkunde, urogynaecologie en het maag-darm-stelsel blijven onderbelicht.

In ***hoofdstuk 4*** presenteren we de resultaten van een systematische review naar de informatiebehoeften en het bijbehorende zoekgedrag van aios rondom vraagstellingen in de klinische praktijk. De 35 geselecteerde studies zijn zeer verschillend opgezet. Er is grote verscheidenheid in de gebruikte methode, de gekozen setting, de selectie van en het aantal deelnemers, de gebruikte meetinstrumenten en de uitkomstmaten. De review laat zien dat zowel huisartsen als specialisten in opleiding in de dagelijkse praktijk gemiddeld één vraag per patiënt hebben, op 83% van de vragen een antwoord proberen te vinden waar ze in 79% van de keren in slagen. Meestal (25-66%) worden collega's hierbij als bron van informatie gebruikt. Aios zoeken vaak tijdens het consult naar een antwoord op hun vraag, al gebeurt dit vaak niet in het bijzijn van de patiënt. Huisartsen in opleiding doen dit vaker tijdens een consult dan specialisten in opleiding. De meeste vragen uit de dagelijkse praktijk betreffen therapie. Aios hebben het idee dat de gevonden antwoorden regelmatig van invloed zijn op het beleid bij een patiënt.

In ***hoofdstuk 5*** evalueren we in een beschrijvende studie onder 76 aios huisartsgeneeskunde hoe vaak ze bij vragen in de dagelijkse klinische praktijk zoeken naar antwoorden en op welke manier zij dit doen. We laten zien dat de informatiebehoeften en het zoekgedrag van derdejaars aios huisartsgeneeskunde in Nederland overeenkomen met resultaten uit eerdere (internationale) studies; ze zoeken in 80% van de vragen en vinden in 80% van de keren het antwoord. Meestal duurt het zoeken vier minuten. In vergelijking met internationale gegevens is het aantal vragen, met een per 5 consulten, echter relatief laag. Aios huisartsgeneeskunde geven aan dat de gevonden informatie in 26% van de gevallen tot een verbetering van de klinische besluitvorming leidt. Naast collega's (28%) gebruiken zij vaak de NHG-Standaarden (26%) als primaire bron van informatie. Andere (online) wetenschappelijke informatiebronnen en originele artikelen waaronder reviews en meta-analyses gebruiken zij weinig.

In ***hoofdstuk 6*** wordt de ontwikkeling en validatie van de Utrechtse EBM kennisvragenlijst beschreven, de "Utrecht questionnaire on knowledge on Clinical epidemiology for Evidence-based Practice" (U-CEP). De U-CEP meet kennis van klinische epidemiologie die nodig is voor de toepassing van EBM. De U-CEP bestaat uit twee vragenlijsten van ieder 25 vragen. In dit hoofdstuk laten we zien dat de U-CEP een valide en betrouwbare kennisvragenlijst is, die kan worden gebruikt om kennis over klinische epidemiologie te meten onder verschillende medische professionals. De U-CEP richt zich hierbij primair op epidemiologische kennis die relevant is voor de dagelijkse klinische praktijk.

Hoofdstuk 7 beschrijft de ontwikkeling en validatie van een instrument dat we ontwikkelden om te meten in welke mate aios huisartsgeneeskunde in de dagelijkse huisartspraktijk

handelen volgens de NHG-Standaarden of hier bewust en onderbouwd van afwijken. Het instrument maakt het mogelijk om drie relevante aspecten van de klinische besluitvorming (diagnose, therapie en verwijzing) van een groot aantal aandoeningen in de dagelijkse huisartspraktijk te meten. Evaluatie onder 12,106 consulten van 76 aios over 12,587 verschillende redenen voor het consult laat zien dat het instrument valide, betrouwbaar en toepasbaar is. Het instrument kan worden gebruikt om het klinisch redeneren en de toepassing van EBM tijdens de opleiding tot huisarts te monitoren.

In ***hoofdstuk 8*** rapporteren we de resultaten van een RCT die we hebben uitgevoerd onder 79 derdejaars aios huisartsgeneeskunde, waarin we de reguliere manier om EBM te onderwijzen hebben vergeleken met een nieuwe, in de klinische praktijk geïntegreerde manier. Bij het reguliere EBM-onderwijs kregen 40 aios huisartsgeneeskunde het EBM-onderwijs alleen op de terugkomdagen, bij het nieuwe EBM-onderwijs kregen 39 aios het EBM-onderwijs deels ook in de huisartspraktijk. Onze RCT laat zien dat met geïntegreerd EBM-onderwijs aios huisartsgeneeskunde EBM niet beter gaan toepassen in de dagelijkse praktijk. Ook de attitude jegens EBM en de kennis over EBM is niet anders dan wanneer ze regulier EBM-onderwijs volgen, niet aan het einde van de opleiding en niet een jaar na het afronden van de opleiding. Met beide vormen van EBM onderwijs handelen aios huisartsgeneeskunde in hoge mate (69-95%) volgens de NHG-Standaarden (of wijken hier bewust van af). Ook hebben ze hetzelfde aantal succesvolle (78-84%) zoekacties (60-82%) bij klinische vragen (14-25%) in de dagelijkse praktijk.

Er wordt, kortom, binnen de huisartsgeneeskunde veel onderzoek gedaan, dat vaak van goede kwaliteit is. De onderzochte thema's sluiten echter nog niet goed aan op wat huisartsen in de praktijk zien. Aios zoeken regelmatig naar antwoorden op klinische vragen, maar vragen dan vooral collega's of zoeken in de richtlijnen; naar andere evidence wordt weinig gezocht. Meer inzicht in de manier waarop EBM wordt toegepast in de praktijk kan helpen het EBM-onderwijs (verder) te verbeteren en meer bij die praktijk te laten aansluiten.

Referenties

1 Sackett DL, Rosenberg WMC, Gray J a M, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. 1996. BMJ. 1996 Feb;312(71).

DANKWOORD

Laat me raden….dit is het enige onderdeel van het proefschrift dat u echt leest?! Zal ik dan eens iets bekennen….ik ook altijd. Om die reden gebruik ik deze ruimte om u toch ook naar de samenvatting te verwijzen. Deze is echt de moeite waard, ook als u niet in de stof zit, niet medisch onderlegd bent, geen huisarts bent of niets van EBM weet. Echt!

Dan echt het dankwoord, misschien wel het belangrijkste hoofdstuk van het hele proefschrift voor mij. Want hoewel ik het toch echt allemaal zelf heb gedaan en ik mijn werklust, doorzettingsvermogen en zelfdiscipline eeuwig dankbaar ben, zou ik het niet hebben kunnen doen zonder hulp.

Mijn grote dank voor ***iedereen die op enige wijze heeft bijgedragen aan het voltooien van dit proefschrift****, waarbij ik een aantal mensen in het bijzonder wil bedanken. Als ik iemand hierbij ben vergeten, sorry (!); promoveren naast opleiding, praktijk en gezin is leuk, fantastisch, leerzaam, uitdagend en nog zoveel meer, maar ook heel vermoeiend.*

Allereerst dank aan mijn promotoren en co-promoren prof. dr. N.J. de Wit, prof. dr. A.W. Hoes, dr. M.E.L. Bartelink, Niek, Arno en Marie-Louise. Mijn dank is groot voor jullie vertrouwen in mij en jullie prettige begeleiding.

Niek, dank dat je mij de kans hebt geboden om te mogen promoveren op jouw bruidsschat. Jouw kritische blik en opbouwende commentaren hebben enorm bijgedragen aan dit proefschrift. Jouw zakelijke, nuchtere visie op zaken bracht mij soms weer terug met beiden benen op aarde. Als ik in mijn enthousiasme weer allemaal extra dingen bedacht om te onderzoeken of te benoemen in artikelen, was jij het die het weer terug kon brengen tot de kern. Jouw scherpe bewaking van de rode draad en jouw kwaliteit om het einddoel in zicht te houden, hebben ervoor gezorgd dat dit proefschrift een geheel is geworden. Dank daarvoor.

Arno, mijn dank en respect voor jouw expertise zijn groot. De manier waarop jij met enkele korte commentaren tijdens een overleg de kern van het probleem weet te raken en daarmee bepalend kan zijn in het proces is onnavolgbaar. Met complimenten ben je niet scheutig, wat het voor mij des te waardevoller maakte als je er een maakte; je hebt geen idee hoeveel energie ik daar weer uit putte. Al liet je me regelmatig lang in spanning zitten voor revisies, ze waren het altijd waard om op te wachten en tilden een manuscript altijd naar een hoger plan. Dank.

Marie-Louise, wat heb ik het getroffen met jou als co-promotor! Jouw flexibele, laagdrempelige, benaderbare manier van begeleiden boden mij veel vrijheid en zorgden ervoor dat ik de PINET studie echt tot mijn studie heb kunnen maken. Jouw enthousiasme voor EBM en jouw plezier in lesgeven werken aanstekelijk. Jouw steun in moeilijkere tijden en oprechte, grote vreugde toen ik zwanger bleek hebben veel voor mij betekend. Naast een fantastisch begeleider, ben je een fantastisch mens. Dank daarvoor!

De leden van de **beoordelingscommissie** wil ik bedanken voor hun inzet, tijd en kritische blik bij het beoordelen van dit proefschrift, prof. dr. R. Jones, prof. dr. R.A.M.J. Damoiseaux, prof. dr. R.J.P.M. Scholten, prof. dr. T.J. ten Cate en prof. dr. N. van Dijk. Ik kijk met een goed gevoel uit naar de verdediging.

Veel dank aan iedereen die heeft bijgedragen aan de dataverzameling voor lopend onderzoek binnen de huisartsgeneeskunde in Nederland; alle leden van het **IOH-r**, alle **hoofden van de verschillende huisartsafdelingen in Nederland, het Nivel en IQ healthcare** en uiteraard **alle onderzoekers** die de vragenlijst(en) hebben ingevuld. Veel dank aan alle **CWO-leden** die in het begin van het traject hebben bijgedragen aan het meedenken over dit onderzoek. Dank uiteraard ook aan het **NHG**, in het bijzonder aan **Jako, Debby en Jacintha**. Dankzij jullie staat de huidige online database over lopend onderzoek binnen de huisartsgeneeskunde als een huis. Wat een mooie prestatie is dat toch! Debby, dank voor de fijne gesprekken altijd weer. Mijn respect voor jouw schoenkeuze is groot.

Dank aan alle **experts** die hebben bijgedragen aan de ontwikkeling van het instrument om het handelen volgens richtlijnen in de huisartspraktijk onder aios te kunnen meten; Guy Rutten, Alfred Sachs, Theo Verheij, Roger Damoiseaux, Lidewij Broekhuizen, Frans Rutten en Kees Gorter. Ook dank aan iedereen die de (voorlopers) van de EBM-kennisvragenlijst (U-CEP) hebben ingevuld, alle **promovendi, aios, experts en opleiders**. Dankzij jullie bestaat er nu een goede, gevalideerde, bruikbare EBM-kennisvragenlijst die gebruikt kan worden in de opleiding.

Veel dank aan iedereen die zijn steentje heeft bijgedragen aan de dataverzameling van de PINET-studie. Alle (inmiddels oud-)**aios uit het 3e jaar van de huisartsopleiding** in Utrecht in 2011; Mijn dank is groot! Dank voor jullie inzet en flexibiliteit. Zeker nu ik zelf in het derde jaar zit en besef dat het zeker niet niets was wat ik van jullie vroeg. Zonder jullie inzet zou dit proefschrift er nooit zijn gekomen. Veel dank ook aan **alle opleiders** van deze (oud-)aios. Dank voor jullie flexibiliteit en medewerking. Dank ook aan alle betrokken **groepsbegeleiders**, **Dorien, Lijan, Michiel, Mariëlle, Mariska, Rianne, Majorie, Peter, Ernest, José, Ellen, Marjam** voor jullie hulp en medewerking. Dan het **managementteam** van destijds, **Monique, Margit, Ron en Rien**, dank voor jullie steun vanuit het management bij het begin van de PINET-studie. De studie werd niet altijd even enthousiast ontvangen; jullie steun en geloof erin waren onmisbaar om het tot een goed resultaat te brengen. Dank! **Aad**, dank voor je tijd en het meedenken over de handigste manier om alle PINET-onderdelen in te plannen in het rooster. Misschien heb ik je toch (een beetje) enthousiast kunnen maken voor EBM? **Willem**, dank voor al je hulp bij ELO. Dankzij jou leerde ik door de bomen het bos zien op ELO. **Marga**, waar zou ik zijn geweest zonder jou? Altijd als ik ergens mee zat of iets wilde bereiken op de huisartsopleiding, kon ik bij jou terecht. Jij wist als geen ander hoe het werkte en waar ik het handigst dingen kon regelen of bij wie ik moest zijn om het geregeld te krijgen. Dank ook voor je luisterend oor en medeleven in tijden

dat het me soms even te veel werd. Jouw oprechte blijdschap op mooie momenten de afgelopen jaren deden me erg goed. **Liesbeth**, net als Marga ben jij onmisbaar geweest in mijn promotietraject en onmisbaar voor alle Utrechtse aiotho's. Ik weet niet of je ooit meer trajectwijzigingen van één aiotho hebt ontvangen, maar je verwerkte ze steeds zonder mopperen en dacht altijd met me mee; veel dank daarvoor!

Dank ook aan alle overige medewerkers van de huisartsopleiding Utrecht voor jullie steun, lieve (opbeurende) woorden, praktische bijstand of hulp bij het uitvoeren van de PINET-studie, het volgen van mijn eigen opleiding aan het instituut of als collega bij het geven van onderwijs. Hoewel ik niet iedereen persoonlijk kan bedanken, wil ik in ieder geval een aantal mensen in het bijzonder bedanken. Allereerst dank aan mijn **groepsbegeleiders,** Lenny en Mariëlle, Margreet en Erik, Maartje en Gerard en Mariëlle en Lijan**.** Ondanks dat ik mijn dubbele rollen soms erg verwarrend vond ("ben ik nou aios, EBM-docent, onderzoeker….?") hebben jullie mij hierin altijd erg prettig bijgestaan en gesteund. Niet altijd was het even makkelijk om alles te combineren en alle ballen in de lucht te houden. Dank voor jullie wakende oog op die momenten dat het nodig was en het vertrouwen dat het goed zou komen. **Margit,** dank voor alle gesprekken over het onderwijs (mede vanuit mijn rol als voorzitter van de landelijke werkgroep Onderwijs van de LOVAH), het sparren over promoveren en het delen van frustraties over het reviseren en publiceren van artikelen. Ik heb deze gesprekken altijd erg gewaardeerd. **Ariëtte,** wat fijn om af en toe gewoon heerlijk je hart te kunnen luchten. Onze gesprekken waren een feest (?) van herkenning voor mij. Dank! **Carolien**, dank voor alle fijne gesprekken over alle het pas-verkregen moederschap en het ongegeneerd mogen showen van foto's. **Charles**, dank voor het sparren over EBM-onderwijs en de samenwerking hierin op de (opleiders)terugkomdagen. Op de workshop op de WONCA hebben we onze ideeën een mooi podium kunnen geven. Wie weet wat er allemaal van komt!

Binnen het Expertisecentrum voor Onderwijs en Opleiding veel dank aan **Lisette en Eugène** voor jullie inbreng en expertise bij de ontwikkeling van de U-CEP. Binnen de afdeling Datamanagement van het Julius Centrum veel dank aan **Susan** voor de verwerking van alle PINET-data.

Dank aan **alle secretaresses** met wie ik in de loop der jaren heb samengewerkt. In het bijzonder wil ik hier **Monique en Jinke** bedanken. Waar zou ik zijn geweest zonder jullie? Of moet ik zeggen, "waar zouden respectievelijk Arno en Niek zijn zonder jullie?" Jullie bewaken en beheersen hun agenda's als leeuwinnen, maar zijn ook nooit te beroerd om mee te denken en te zoeken naar gaatjes en mogelijkheden. Veel dank hiervoor! **Coby**, dank voor het wegwijs maken door het Julius en meedenken bij het zoeken naar ruimtes om te vergaderen (als deze er eigenlijk niet meer waren).

Dank aan alle **co-auteurs en reviewers** van de artikelen in dit proefschrift voor jullie nuttige commentaren en revisies. In het bijzonder veel dank aan **Peter**; dank voor je eindeloze geduld, uitleg, flexibiliteit en hulp bij de analyses van de PINET-studie.

Zonder jouw kennis en inzicht in cluster-gerandomiseerde trials zou ik mijn analyses nooit hebben kunnen volbrengen. Veel dank! **Alma,** veel dank voor al je hulp bij het opzetten, uitvoeren en opschrijven van de studie naar lopend onderzoek binnen de huisartsgeneeskunde. Jouw kritische opmerkingen en scherpe analyses hebben mij enorm geholpen bij het schrijven van mijn eerste artikel als eerste auteur. Dank! **Geert**, dank voor je verfrissende inzichten bij de verschillende onderdelen waaraan je hebt bijgedragen. Vaak bekeek jij zaken net vanuit een andere invalshoek waardoor het geheel weer beter kon worden. Dank daarvoor! **Esther,** dank voor het delen van je schat aan kennis over onderwijs binnen de medische (vervolg)opleidingen. Ik ken niemand die zo goed in de literatuur zit als jij. Geweldig!

Veel dank ook aan alle (destijds) studenten die hebben geholpen bij het verzamelen en verwerken van de data betreffende het onderzoek binnen de huisartsgeneeskunde. **Alicia**, nog voor ik officieel aan mijn aiothotraject was begonnen hielp jij me als eerste student "onder mijn hoede" bij het verzamelen van de data voor hoofdstuk 2 over de RCT's binnen de huisartsgeneeskunde. Je doorliep alle titels en abstracts alsof het je lust en je leven waren en had alles binnen een mum van tijd gepiept. Dank! **Lisanne,** als student hielp jij me bij het in kaart brengen van al het lopend onderzoek in Nederland binnen de huisartsgeneeskunde. Dank voor het afstruinen van websites op zoek naar aanknopingspunten om dit in kaart te kunnen brengen en te categoriseren.

Veel dank ook aan alle (destijds) studenten en (inmiddels) mede-aios huisartsgeneeskunde die hebben geholpen bij het verwerken van de data van de PINET-studie; Imke, Lisanne, Yvonne en Jeroen. Dank voor jullie vele uren ploeteren door de database en eerste analyses van de data met betrekking tot het handelen van richtlijnen door aios. Super dat jullie (ondanks dat?) allemaal enthousiast zijn gebleven voor het huisartsenvak en allemaal de huisartsopleiding zijn gaan doen. **Imke,** veel dank voor jouw tomeloze energie en werklust bij het ontwikkelen van het instrument om het handelen volgens richtlijnen door aios huisartsgeneeskunde te kunnen meten. Jij hebt de solide basis gelegd voor het uiteindelijke instrument. Mijn dank is groot! **Lisanne**, wat heb jij je vastgebeten in het onderwerp. Jij hebt het instrument verfijnd en de spits afgebeten bij het scoren van de ruim 3000 patiëntconsulten. De vele uren werk hebben je niet afgeschrikt gelukkig; je bent nu zelfs aan je eigen aiothotraject begonnen. Superleuk dat je verder gaat waar ik ben gebleven! **Yvonne**, als co-assistent in mijn eerstejaars praktijk uitte je interesse in onderzoek doen binnen de huisartsgeneeskunde en dat heb ik met beide handen aangegrepen. Daar heb ik geen moment spijt van gehad! Mede dankzij het voorwerk van Lisanne leek het je allemaal gemakkelijk af te gaan en ploeterde je je met gemak door de ruim 3000 consulten heen. En passant hielp je me zelfs nog verder op weg door Nicole aan te bevelen, die ook graag onderzoek wilde doen. Dank voor alles! **Jeroen**, jou heb ik enthousiast kunnen krijgen tijdens jouw START-blok, waarin ik jouw groep begeleidde bij het schrijven van jullie eerste (zeer succesvolle) EBCR. Als enige mannelijke student

(of moet ik zeggen "KNO-arts"?) hielp jij me bij de laatste fase van de dataverwerking. Jouw stage begon vlak nadat ik voor het eerst moeder was geworden. Je luisterde naar al mijn verhalen hierover en deed net alsof je niet anders gewend was. Ook in het onderzoek leek het alsof je dit al jaren deed; wat een feest zo vlak na mijn verlof! Het was me een eer en groot genoegen om jouw afstudeerpraatje te mogen houden.

Lia, wat een feest dat jij in je derde jaar als differentiatie wetenschap koos. Dankzij jou kregen we inzicht in het zoekgedrag van (ons als) aios huisartsgeneeskunde in de dagelijkse praktijk. Klap op de vuurpijl was onze duo-presentatie op de WONCA! Dank voor je inzet, flexibiliteit en doortastendheid. **Nicole,** na aanbeveling door Yvonne kwam je mij helpen bij het uitvoeren van de review zoals beschreven in hoofdstuk 4. In je vrije tijd (!) screende je naast je werk vele titels en abstracts en hielp zo bij de selectie van de artikelen. Dank voor je vele uren inzet en bijdrage aan dit proefschrift. **Helen,** getipt door Esther kwam je mij helpen bij de analyses voor de validatie van de U-CEP, "gewoon" omdat je dit leuk vond om te doen. Veel dank hiervoor!

Dank aan mijn (vele) **kamergenoten en mede-promovendi** in de loop der jaren, Henri, Douwe, Mirjam, Karlijn, Christiana, Fleur, Martine, Pushpa, Minke, Laura, Hadassa, Marjolein, Mariëtte, Wouter, Rose, Eva en Myrthe (en eenieder die ik hier (toch) vergeten ben). Dank voor de gezellige kameruitjes, jullie luisterend oor en jullie medeleven op mooie en verdrietige momenten. In het bijzonder veel dank aan **Henri**, voor het creëren van mijn uitzicht op de klaproos, al je hulp bij R en je geduld hierbij. Altijd weer wist jij me met jouw analytisch vermogen en slimme, snelle syntaxen verder op weg te helpen als ik (weer eens) was vastgelopen. Jouw statistische kennis en onderzoekskwaliteiten zijn bijzonder ("je snapt het gewoon ècht…!"); ik heb er alle vertrouwen in dat jij het ver gaat schoppen binnen de onderzoekswereld! **Douwe**, dank voor de fijne gesprekken over het ziekenhuis, de beslommeringen als aios en het leven als vader/moeder. Mooi om te zien hoe je het allemaal doet en er zo van kan genieten. **Mirjam**, dank voor de vele potten thee, ook namens de plant(en) op onze kamer. Dankzij jou hadden wij zowaar een echt levende plant, weet ik "alles" over wedstrijddansen en heb ik geleerd hoe je van je auto een kledingkast kan maken. Dank! **Karlijn**, excel- en syntaxkoningin! Jij hielp me bij het terugvinden van data met snelle, handige, slimme syntaxen en hebt er, net als Henri, voor gezorgd dat ook ik uit de voeten kon met R. Mijn dank en respect voor jouw kennis zijn groot!

Veel dank aan mijn **opleiders**, Mieke, Wietze, Koen, Roeland en Dorine. Dankzij jullie was ik in staat om naast de praktijk en mijn opleiding te werken aan mijn promotie. **Mieke**, mijn tijd bij jou in de praktijk was het zwaarste jaar ooit, al kwam dat zeker niet door jou of de praktijk! Je verwoordde het laatst nog zo mooi in één zin die de hele lading dekte . Daarmee werd het voor mij weer opnieuw duidelijk hoe zeer je hebt meegeleefd; dank daarvoor! **Wietze**, dank voor het zijn van zo'n goed voorbeeld. De manier waarop jij doktert is hoe ik het ooit hoop te kunnen. Fijn dat ik mijn twijfels met jou kon bespreken. Het is mede dankzij jou dat ik bij mijn keuze gebleven ben.

Dank daarvoor! **Koen**, veel dank voor je nuchtere blik op zaken. Ik kijk met een goed gevoel terug op onze leergesprekken, wandelend aan het water in het zonnetje, met discussies over zoveel meer dan alleen het huisartsenvak. Mijn steeds groter wordende buik heeft er gelukkig (bijna) nooit voor gezorgd dat dit niet meer ging. Dank voor de fijne tijd in jouw praktijk! **Roeland**, dank voor je begrip zo in de eerste weken in de praktijk, vlak na m'n verlof waarin ik ook de leesversie moest afmaken. Ik stond echt op 'standje overleven', maar het is gelukt. Nu eindelijk weer tijd en ruimte om mijn laatste resterende tijd als aios bij jou in de praktijk volop te benutten. Ik heb er zin in! **Dorine**, dank dat je zo met me meedenkt en er mede voor waakt dat alle ballen in de lucht niet te veel worden. Ik kijk uit naar nog een halfjaar, waarin er in ieder geval weer een bal minder in de lucht gehouden moet worden. Naast al mijn opleiders ook veel dank aan alle **assistentes en andere huisartsen** in de praktijken voor jullie interesse, medeleven en vaak interessante vragen over mijn promotie.

Dank aan alle **collega-aios** uit de verschillende groepjes waar ik in de loop der jaren in heb gezeten voor jullie interesse in mijn promotie en jullie steun op momenten dat het soms even nodig was. In het bijzonder veel dank aan **Larissa**. Dank voor je gezelligheid, je oneindige interesse en je vriendschap. Al zien we elkaar niet vaak, het voelt voor mij altijd alsof ik je de dag ervoor nog heb gezien. Dank ook dat je me tijdens de WONCA (als schone slaapster) gezelschap hield als ik zat te ploeteren aan de scriptie voor mijn master. Leden van mijn supervisiegroepje, **Louise, Pieter, Lara en Janneke**, dank voor jullie open houding en kritische vragen tijdens de bijeenkomsten. Dankzij jullie lukt het me nog beter om te prioriteren en te (her)waarderen wat echt belangrijk is in het leven.

Dank aan mijn twee voorbeelden op onderzoeksgebied, Peppino en Patrick. **Peppino**, door jouw relaxte houding over onderzoek en jouw doortastendheid om dingen "even" te regelen tijdens mijn zesdejaars-stage heb je me (onbewust) enthousiast gemaakt om onderzoek te (gaan) doen. Door jou dacht ik "hé, misschien kan ik dat ook wel", al kwam dat pas jaren later na wat omwegen tot uiting. Dank! **Patrick**, wat had ik graag gehad dat jij er nog was en er gewoon bij had kunnen zijn op mijn promotie om alsnog "dat biertje" te doen, 10 jaar nadat mijn onderzoek klaar was en we de 1e opzet voor het artikel af hadden. Ik denk nog vaak terug aan jouw afstudeerpraatje destijds en jouw gestelde vertrouwen in mij. Je hebt geen idee gehad hoeveel dat voor mij betekende. Dank daarvoor!

Mijn dank ook aan de **SBOH**, voor het mogelijk maken om als aios huisartsgeneeskunde de opleiding te combineren met een promotietraject. Dankzij jullie financiële steun kon ik menig buitenlands congres bezoeken en onze (nieuwe) inzichten over EBM-onderwijs verspreiden. Ik heb het gevoel dat het EBM-onderwijs mede dankzij deze steun weer meer op de huisartsgeneeskundige kaart is gekomen, zowel nationaal als internationaal. Ook dank aan het bestuur van de **LOVAH**, het aiothobestuur en de werkgroep onderwijs; veel dank voor jullie inzet om het onderwijs en het aiothoschap verder te verbeteren.

Hieronder veel dank aan iedereen die mijn leven zoveel fijner, leuker of gezelliger maakt. Het is deze basis die ervoor zorgt dat het me is gelukt dit proefschrift af te ronden en mij de ontspanning biedt die hierbij nodig is. Dank!

Lieve dames van "**Factor 8**", lieve Caro, Fleur, Ka, Lon, Lies, Lau, Noor, Pien en Vi('tje): wat een feest dat wij in het derde jaar van geneeskunde bij elkaar in de groep kwamen! Ik geniet nog altijd enorm van onze etentjes en avondjes. Met z'n allen kunnen we ongegeneerd alles bespreken waar we in ons werk - al dan niet in combinatie met privé – tegenaan lopen. Nu de eersten van ons aios-af zijn en door mijn promotie meer dan de helft (!) van ons is gepromoveerd, durf ik met recht te zeggen dat wij stiekem toch echt allemaal wel een beetje nerds zijn. Ik ben trots op jullie allemaal en trots dat we nog steeds, na al die jaren, zo hecht zijn. Dank voor jullie vriendschap!

Lieve **vrienden en vriendinnen**, dank voor jullie gezelligheid, luisterend oor en interesse in mijn promotie. Lieve **Aukje,** dank voor de avondjes gezelligheid onder het genot van een – door jou gekookt – maal. Dank ook voor je lieve woorden als ik ze nodig had. Ik ben blij dat we nu allebei met twee koters bij elkaar op de koffie kunnen komen. Gelukkig heb ik daar nu (eindelijk) weer meer tijd voor! Lieve **Marieke**, dank voor je kaartjes, smsjes en belletjes. Op de een of andere manier weet jij altijd precies aan te voelen wanneer deze nodig zijn; dank daarvoor! Lieve **Coennie**, wat zit je ver weg……. Gelukkig is er de telefoon en al spreken we elkaar niet vaak, als we elkaar spreken is het net alsof je gewoon nog bij me om de hoek woont. Ik ben blij om te zien dat het zo goed met je gaat en dat je het zo enorm naar je zin hebt in "Verweggistan". Lieve **Eef**, wat een feest om jouw kamergenoot te mogen worden aan het begin van mijn promotietraject! Jouw enorme hoeveelheid energie en levenslust werken aanstekelijk. Je bent een topwijf! Lieve **Kate**, wat ben ik blij dat wij na al die jaren nog steeds vriendinnen zijn! Dank voor alle mooie avondjes en mooie momenten. Dank dat ik bij jou altijd mezelf kan zijn en dank voor je onbreekbare vriendschap. Ik ben enorm trots hoe jij en Erik het de afgelopen jaren hebben gedaan en hoe jullie alle tegenslagen een plek hebben kunnen geven. Je hebt weleens gezegd dat ik in sommige dingen een voorbeeld voor je ben; dat is vice versa net zo! Lieve **Eline**, samen opgroeien, samen hockeyen, samen in een studentenhuis wonen, en (uiteindelijk) ook samen onze kinderen zien opgroeien. Wat kan ik genieten als ik onze mannen samen zie spelen; hoe bijzonder is dat! Wat ben ik trots en blij dat wij nog steeds vriendinnen zijn. Dank voor onze vriendschap! Lieve **Fleur**, naast alle factor-8-etentjes zijn wij in de loop der jaren steeds meer samen met onze mannen erbij gaan afspreken om onder het genot van lekker eten en goede wijnen mooie gesprekken te hebben. Deze avondjes zijn mij (en Pieter) veel waard. Gelukkig zijn Roeland, jij en Quint pas geleden verhuisd naar een werkelijk prachtig huis met briljante eetkeuken, dus dat moet goed komen. Ik kijk uit naar nog vele mooie avonden! Lieve **Lon**, vriendin uit duizenden. Waar moet ik beginnen met bedanken? Dank voor alle mooie avondjes en goede gesprekken. Al woon je niet om de hoek en zien we elkaar niet dagelijks, de momenten dat we elkaar zien zijn altijd fijn, prettig en

waardevol. Nu met de uitbreidingen van onze gezinnen wordt de aandacht iets verlegd van mooie avondjes met veel te veel lekker eten en lekkere wijnen naar mooie dagen vele prachtige momenten met de kids. Ik kijk uit naar alle komende momenten om te delen en verheug me nu alweer op Oud&Nieuw, met of zonder alcohol ;-) Dank voor wie je bent! Lieve **Bart**, als enige man in deze lijst van vrienden en vriendinnen. Dank dat ik je al ruim 30 jaar (!) mijn vriend mag noemen. Dank voor je relativeringsvermogen, je rust en je heldere kijk op zaken. Al vind ik dat je tegenwoordig eigenlijk veel te ver weg woont, ik ben ongelooflijk trots op je wat je daar allemaal voor elkaar krijgt en hoe je het doet. Je bent goud waard! Lieve **vrienden en vriendinnen van Pieter**; dank dat jullie hem af en toe meenamen, zodat ik verder kon werken.

Lieve (oud-) **leden van Stichting Geron**, lieve **Annekoos, Bastiaan, Maria, Henk en Taco**; dank voor alle mooie ervaringen die ik met en dankzij jullie heb kunnen opdoen in en over India. We delen de liefde voor dit prachtige land en ik hoop dat wij hier nog lang mee door kunnen gaan.

Lieve **Angelique**, dank voor al je hulp in ons huis, maar vooral ook dank voor je luisterend oor en kritische vragen en opmerkingen over verschillende zaken. Zonder jouw hulp was het me nooit gelukt om het huis nog enigszins leefbaar te houden in de laatste maanden van mijn promotie. Dank!

Lieve **schoonfamilie**, dank voor de warme ontvangst in de familie. Al snel voelde ik me thuis. De jaarlijkse wintersportvakanties met "het grote gezin" zijn goud waard. Behalve skiën heb ik hier nog veel meer geleerd; dat artsen, burgemeesters en ergonomen ook maar "gewoon mensen" zijn. Dit inzicht relativeerde veel in mijn leven en heeft zeker bijgedragen aan de manier waarop ik nu in het leven sta. Dank daarvoor! Lieve opa's en oma van Pieter, lieve **opa Chef en oma Poes**, dank voor jullie interesse in mijn opleiding en promotie. Ik had graag gehad dat jullie er vandaag bij konden zijn, maar weet zeker dat jullie van boven meeleven en meegenieten. Jullie zijn een voorbeeld voor mij (geweest)! Lieve **opa Piet,** dank voor je lieve woorden op ons huwelijk. Je ziet waar het toe kan leiden; het heeft het (zelfs) geschopt tot een van mijn stellingen: "Wat je niet begrijpt, moet je bewonderen". Dank voor deze wijze levensles! Lieve **Ly en Eef**, dank voor jullie voorbeeld, als enige gepromoveerden in mijn (schoon)familie. Jullie hebben regelmatig aangegeven niet alleen maar jaloers te zijn dat jullie promotie al ver achter jullie ligt. Misschien juist hierdoor wel dat ik het des te meer (ook) wilde? Dank ook voor de fijne gesprekken over het dokterschap en het veranderende klimaat om als arts in te werken. Jullie hebben geen idee hoeveel waarde ik hecht aan jullie ervaringen en mening hierin. Dank! Lieve **Hanna en Jap**, dank voor de deur die altijd open staat. Ik geniet van de – meestal ongeplande – bezoekjes aan jullie in jullie heerlijke huis. Jullie kwaliteiten om van de meest normale dingen iets bijzonders te maken is ongekend. De zelfgebakken scones, de vers gemaakte koffie en de uit de losse pols geklopte chocoladetaartjes zijn toch stiekem

wel hoogtepuntjes om naar uit te kijken. Dank ook voor jullie begrip en steun in tijden dat het allemaal wat zwaarder was. Jullie zijn schatten! Lieve **Goof**, dank voor je nuchtere no-nonsense visie op het leven. Hoewel de keuzes die jij maakt in het leven soms mijlenver afstaan van de mijne, lijken we misschien in sommige opzichten toch meer op elkaar dan je op het eerste gezicht zou denken. Ik heb veel respect voor de manier waarop jij bent omgegaan met tegenslagen. Je hebt ongekende kwaliteiten om het iemand naar de zin te maken. Zolang je jezelf daarbij niet uit het oog verliest is dat jouw kracht! Ik geniet ervan op de wintersport. Dank! Lieve **Ernst en Annelies**, of moet ik zeggen "lieve opa Ei en oma Lies". Allereerst; dank voor het grootbrengen van zo'n fijne zoon en het zijn van zo'n fijne opa en oma voor Koene en Kate. Mede hierdoor liet ik ze met een gerust hart achter en kon ik, zeker in de laatste maanden, werken aan mijn promotie. Dank ook voor jullie hulp bij grote keuzes in ons leven, in raad en daad. Het vangnet dat jullie hebben gecreëerd biedt Pieter en mij kansen en mogelijkheden om onze dromen te verwezenlijken. Dank voor jullie onvoorwaardelijke steun en vertrouwen in ons. Lieve **Kaas en Relinde**, wat een feest dat jullie er misschien helemaal niet bij zijn op mijn promotie! Ik had me geen betere reden kunnen bedenken. Jullie zijn als geen ander in staat om beren op de weg te knuffelen. Wat heb ik een respect voor de manier waarop jullie in het leven staan en bewust en weloverwogen jullie keuzes maken. Jullie vermogen om in oplossingen te denken in plaats van in problemen opent vele deuren voor jullie en heeft er voor gezorgd dat jullie straks/nu (?) papa en mama zijn. Het is jullie gegund!

Lieve **paranimfen**, lieve Lobke en Inge; wat een eer dat jullie naast mij staan tijdens mijn promotie. Jullie zijn beiden als geen ander in staat met beide benen stevig op de grond te staan en te relativeren. En dat is precies wat ik bij mijn promotie wel kan gebruiken! Lieve **Lob**, wie anders dan jou had ik naast me willen hebben op mijn promotie? Als getuige bij jouw huwelijk en paranimf bij jouw promotie kon ik niet wachten je te vragen voor deze rol. Tijdens het maken van de almanak "Extase" is de solide basis voor onze vriendschap gelegd. Samen met jou ben ik vervolgens naar India geweest, een reis vol mooie en bijzondere ervaringen. Daarna is onze vriendschap alleen maar verder versterkt en hebben we vele mooie avonden en etentjes beleefd met onze mannen. Wat zijn Siem en jij toch fantastisch in staat om jullie gasten te verwennen! Lieve Lob, jouw enorme doorzettingsvermogen en werklust zijn onnavolgbaar. Als jij een weg hebt uitgestippeld kan zelfs een bulldozer je hier nog niet vanaf helpen. Dat maakt je tot iemand op wie ik altijd kan vertrouwen en bij wie ik altijd terecht kan als dat nodig is. Dank daarvoor! Lieve **Ing**, wie had gedacht dat jij mijn paranimf zou zijn tijdens mijn promotie toen wij vele jaren geleden samen in het MDT-bestuur zaten? Door wederzijdse vriendinnen en hockey zijn onze wegen elkaar (op meerdere manieren) weer gekruist. Ik kan niet anders zeggen dan dat ik daar enorm blij mee ben. Ing, ik ken niemand die zo klaarstaat voor haar vriendinnen als jij. Als jij eenmaal iemand in je hart hebt gesloten, komt diegene daar niet zomaar meer uit. Net als Lob heb jij een

enorme ambitie, werklust en doorzettingsvermogen. Ik bewonder de manier waarop jij je carrièremogelijkheden uitstippelt en daarin hele bewuste keuzes maakt. Ik heb er alle vertrouwen in dat je daar belandt waar je wil belanden.

Lieve **familie**, waar zou ik zijn zonder jullie? Jullie vormen de basis is mijn leven en bieden mij de structuur die ik nodig heb om van mijn vrijheid te kunnen genieten. Jullie onvoorwaardelijke liefde, steun en vertrouwen zijn mij heel veel waard. Lieve **oma Borsboom**, jij bent altijd mijn grote voorbeeld geweest. De band die ik met jou had is niet te beschrijven. Lieve oma, ook al ben je al lang niet meer hier op aarde, toch geloof ik dat je er bij bent bij mijn promotie; dank daarvoor! Lieve **papa en mama**, dank voor jullie onvoorwaardelijke steun en vertrouwen bij alles, altijd. Jullie hebben mij altijd alle kansen gegeven om mij verder te ontwikkelen, op alle mogelijke manieren en in alle mogelijke richtingen. Niets was jullie te gek. Zonder jullie zou ik nooit zo ver zijn gekomen. Dank ook voor jullie steun toen niet alles lukte wat Pieter en ik graag wilden. Jullie medeleven daarbij gaven ons veel steun en vertrouwen in de toekomst. Als Pieter en ik maar half zo goed in staat zijn om zo onvoorwaardelijk van onze kinderen te houden en hen zo onvoorwaardelijk te steunen, wacht hen een gouden toekomst. Dat weet ik zeker. Lieve **zussen,** lieve Lin, Jon en Baab. Wat lijken we op elkaar en wat zijn we verschillend. Het is een voorrecht om jullie zus te mogen zijn en zo mogelijk een nog groter voorrecht te mogen zien hoe jullie terecht zijn gekomen. Lieve **Lin**, als oudste van vier een bijzondere positie binnen het gezin. Ik heb grote bewondering voor de manier waarop jij kan genieten van de kleine dingen van het leven. Ik waardeer jouw vermogen om los te laten en te laten gebeuren. Wat ben ik trots op je en wat geniet ik ervan om te zien hoe jij het voor elkaar hebt met Sander, Rick en Lieke. Lieve **Jon**, mijn eerste grote zus die in een andere stad ging studeren en wonen. Het is dankzij jou dat ik het heel normaal vond om deze stap ook te zetten en te gaan studeren, lid te worden en op kamers te gaan. Dank voor jouw voorbeeld daarin. Ik heb veel respect voor jouw altijd aanwezige gevoel voor rechtvaardigheid. Bij jou is het altijd duidelijk wat iemand aan je heeft en waar je voor staat. Ik ben trots op de grote stappen die je de laatste tijd hebt genomen en heb er alle vertrouwen in dat je daar komt waar je wil zijn. Met Ramses, Carolijn, Matthijs en Marije heb je een prachtgezin! Lieve **Baab**, lief klein zusje. Wat ben ik trots op je. Je hebt belangrijke beslissingen genomen de laatste jaren en het lef gehad je leven om te gooien. Ik geniet er enorm van om te zien hoe blij jij bent om nu (ook) in de zorg te werken. De vele gesprekken over de zorg en patiënten zijn mij veel waard. Hoewel ik met heel mijn hart hoop dat je er bij bent bij de promotie, ben ik erg blij dat ook jij (binnenkort/zojuist) moeder bent. Ik weet zeker dat je een hele goede moeder zult zijn! Dank ook voor het maken van mijn omslag; zonder jou zou mijn proefschrift er nooit zo mooi hebben uitgezien. Een half woord was genoeg om precies te maken wat ik wilde. Dank daarvoor! Lieve **Ramses, Sander en Dirk**, lieve mannen van mijn zussen. Ik ben blij dat jullie onze familie zijn komen versterken. Dankzij jullie is er een andere dynamiek gekomen, wat onze familieband alleen maar heeft versterkt.

Dank! Lieve **Carolijn, Matthijs, Marije, Rick en Lieke**, met jullie komst is onze familie nog verder uitgebreid en versterkt. Ik geniet ervan om jullie te zien opgroeien en ben een immens trotse tante!

Lief **gezin**, lieve Pieter, Koene en Kate. Jullie zijn mijn alles. Hoeveel ik ook van mijn werk houd en geniet, iedere dag kijk ik er weer naar uit om naar huis te kunnen komen en jullie in mijn armen te mogen sluiten. Waar je zou verwachten dat een jong gezin energie vreet en betekent dat het "tropenjaren" zijn, geeft het mij (daarnaast) vooral ook heel veel energie. Dank voor alles, dank dat jullie in mijn leven zijn en dank voor wie jullie zijn. In mijn ogen zijn jullie geweldig! Lieve **Koene**, jouw onuitputtelijke energie, jouw nieuwsgierigheid, jouw wil om alles te ontdekken, jouw enorme doorzettingsvermogen en jouw wilskracht maken het opvoeden van jou tot een groot plezier en een grote, schitterende uitdaging. Ik geniet van alle momenten met jou. Steeds opnieuw verbaas je me met wat je nu weer allemaal kan én wil. Ik voel me enorm bevoorrecht jouw mama te mogen zijn. Ik ben zo ontzettend trots op je en kijk uit naar alle mooie momenten die nog gaan komen. Ik hou van je! Lieve **Kate**, mijn heerlijke, fijne, tevreden, mooie meisje! Jouw prachtige glimlach straalt me 's ochtends al tegemoet als ik jou uit bed kom halen en maakt dat ik iedere dag goed begin! Jouw oprechte woede als iets dat je wil niet lukt en jouw immens tevreden blik als het dan toch is gelukt beloven veel goeds. Ik ben nu al zo ongelooflijk trots op je. Ik kijk er naar uit om jou verder te zien opgroeien. Lieve Kate, ik hou van jou! Lieve **Pieter**, waar zou ik zijn zonder jou? Mijn steun, mijn toeverlaat, mijn vriend, mijn man, vader van onze kinderen en nog zo veel meer! Jouw vermogen om te verwoorden wat ik bedoel en mijn vermogen om te doen wat jij denkt zijn illustratief voor de manier waarop wij samen sterk zijn. Ik bewonder jouw vermogen om last minute dingen op sublieme wijze voor elkaar te krijgen. Daar waar ik 's ochtends vlam, ben jij 's avonds het meest in je element; ook dat maakt ons samen tot een gouden team. Samen zijn wij de uitdaging aangegaan om onze kinderen het plezier in het leven mee te geven wat ons zo dierbaar is. Samen kunnen wij de wereld aan. De uitspraak "Zou je dat nou wel laten?" dekt voor mijn gevoel precies hoe wij in het leven staan. Ik kijk uit naar de rest van ons leven samen en heb zin in alle projecten die wij nog gaan oppakken. Dit project is nu in ieder geval afgesloten; tijd voor een nieuwe. Doe je mee?

Marlous Koningsveld-Kortekaas

CURRICULUM VITAE

Marlous Frédérique Kortekaas was born on the 13th of March 1983 in Wassenaar, the Netherlands, the third of four daughters. After graduating cum laude from the Gymnasium at the 'Sint Adelbert College' in Wassenaar in 2001, she began her medical education at the University of Utrecht. Throughout her medical education Marlous was an active member of various (medical) boards and committees. She completed 2 internships in India and Brazil and combined them both with travel and working as a volunteer for an NGO in India. Focus of the NGO in Ahmednagar, Snehalaya, was to improve the lives of deprived women and children; with HIV/AIDS, who had been sexually abused, were orphans or prostitutes. Since then she has been a board member of 'Stichting Geron', an organization supporting similar projects in India, both financially and through consulting.

Marlous has a strong desire instilled in her to want to improve the quality of healthcare in The Netherlands. After obtaining her medical degree in 2007, she therefore decided to join the Plexus Medical Group as a business analyst to help change the structure and network of the healthcare system. After one year Marlous came to the realisation she wanted to train as a physician so started a two year residency in the Emergency Room at the MESOS Medisch Centrum in Utrecht and the Zuwe Hofpoort Ziekenhuis in Woerden. During this residency Marlous and her husband bought and completely renovated their family home themselves.

In 2010 she reached out to her promoter to initiate a PhD research project that matched her interest; to improve health care in the field of general practice (GP). Once Marlous was admitted to the postgraduate GP specialist programme she combined this education with the research described in this thesis and a PhD at the Julius Center for Health Sciences and Primary Care at the Utrecht Medical Center. In parallel she also successfully completed a Master of Science in Clinical Epidemiology at the University of Utrecht in 2013, frequently gave lectures for the GP specialist curriculum on EBM and (has) almost obtained her Basic Qualification in Education. Throughout her GP specialist programme Marlous was an active member of the national board of AIOTHOs, and chair of the national education committee of the Dutch National Association of GP Trainees (LOVAH) (both 2010-2013). The first unites all Dutch GP trainees who combine their GP specialty training with a PhD project, the second represents all Dutch GP trainees and aims to guard and improve quality of the GP specialist programme.

At present, she is in her last year of the GP specialty training and expects to complete her GP degree in January 2017 to start working as a GP thereafter. She lives with her husband (Pieter) and two children (Koene (2014) and Kate (2015)) in Utrecht.