Directed self-learning as approach to continuing professional development

Citation for published version (APA):

Hobma, S. O. (2005). *Directed self-learning as approach to continuing professional development*. [Doctoral Thesis, Maastricht University]. Datawyse / Universitaire Pers Maastricht. <https://doi.org/10.26481/dis.20050624sh>

Document status and date:

Published: 01/01/2005

DOI:

[10.26481/dis.20050624sh](https://doi.org/10.26481/dis.20050624sh)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Directed self-learning as approach to continuing professional development

THE EFFECT OF A MULTI-LEVEL INTERVENTION

The study described in this thesis was funded by ZonMw, the Netherlands Organization for Health Research and Development (project number 1615.0005).

Lay-out: P.Hinssen, N.J.M. Meijers, Datawyse

Omslag: N.J.M. Meijers, Datawyse

Ets omslag: G.B. Piranesi

ISBN: 90-5278-465-5

© Sjoerd Hobma, Maastricht 2005

Directed self-learning as approach to continuing professional development

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Maastricht,
op gezag van de Rector Magnificus,
Prof. mr. G.P.M.F. Mols
volgens het besluit van het College van Decanen
in het openbaar te verdedigen op
vrijdag 24 juni 2005 om 14.00 uur

door

Sjoerd Hobma

Promotores

Prof. dr. R.P.T.M. Grol

Prof. dr. C.P.M. van der Vleuten

Co-promotor

Dr. P.M. Ram

Beoordelingscommissie

Prof. dr. H.F.P. Hillen (voorzitter)

Prof. dr. J. Heyrman (Katholieke Universiteit Leuven)

Prof. dr. J.M.J.P. van der Linden

Prof. dr. A.J.J.A. Scherpbier

Dr. M.J.P. Wensing (UMC St. Radboud Nijmegen)

CONTENTS

Chapter 1	Introduction	7
	DEVELOPMENT	
Chapter 2	Directed self-learning for medical professionals: development of an intervention to support CPD	13
Chapter 3	GPs' use of assessments for continuing professional development	31
Chapter 4	Setting a standard for performance assessment of doctor-patient communication in general practice	41
	EFFECTIVENESS	
Chapter 5	Effectiveness in improving doctor-patient communication	53
Chapter 6	Effectiveness in improving clinical performance on ENT-disorders	65
	FEASIBILITY	
Chapter 7	Feasibility and appreciation	77
	DISCUSSION	
Chapter 8	General discussion	91
	Summary	107
	Samenvatting	113
	Appendix	119
	Authors and publications	120
	Dankwoord	123
	Curriculum vitae	127

Chapter 1

INTRODUCTION

This thesis describes a study on continuing professional development of general practitioners in The Netherlands. This chapter provides the context and the rationale of the study, the specific research questions, the general study-design and the outline of the thesis.

Rationale of the study

The term ‘continuing professional development’ (CPD) refers to the necessary ongoing education of doctors during their careers. It describes the process of maintaining the competence to practise (learning) and of optimising patient care (improving). At the time of graduation and certification, GPs’ knowledge and skills are supposed to be of a level that allows them to participate in patient care. From then on, they are expected to deliver adequate patient care until the moment of their retirement. This is possible only if knowledge and skills are appropriately used and maintained.

The need for doctors to learn and improve has increased due to the changing context of health care. This can be attributed to a number of factors, such as the expanding medical insights, the need to maximise efficient use of available funding for health care, and the growing demands of patients.

Maintenance of competence and adequate use of knowledge and skills in daily practice are regarded to belong to a doctor’s professional responsibility. Although they should account for their learning and improving, doctors need support in their continuing professional development. During the last decades, the focus of extensive educational research and development has been on how exactly this support must be given. General theories on learning and improvement of professionals, like problem based learning, adult learning theory or the stages of change model are available.^{1 2} In the specific context of medicine, approaches like Evidence Based Medicine have been developed, and have led to new instruments, like guidelines that support physicians in learning and improving care.³ The developed methods and approaches of educational support for doctors have in itself been the object of study, which has also attributed to the insights in how professional development can be supported.⁴⁻⁷

The general conclusion of this educational research is that an agreed ‘best approach’ to support the continuing development of doctors is not available. Not one of the theories, approaches or instruments has yet demonstrated to be superior in supporting physicians in optimising patient care.^{8 9} In literature a number of characteristics that influence the effectiveness of educational interventions have been defined. They must be ‘learner-centred’ and tailored to specific needs of health professionals and their working context. Also, learning activities must be ‘self-directed’ which means that they are based on the perceived needs of the learning professional. Maybe, multifaceted interventions, i.e. educational interventions combining different methods, are more effective than single interventions. These characteristics have been criticised as being based on beliefs and on limited or conflicting evidence.^{7 8 10} Moreover, it is not clear how they can be translated into an educational intervention that supports the CPD of doctors. These are the starting points of our study.

The first aim of the study described in this thesis is to further develop and ground an approach to support general practitioners (GPs) in their continuing professional development on the basis of the available insights. The second aim is to investigate the effectiveness and feasibility of the newly developed approach. Effectiveness means that it results in adequate improvement of delivered care. Feasibility comprises the ability to perform the educational approach in the current context, in terms of available time, money, and educational skills and appreciation of the target group.

Context of the study

GPs in The Netherlands are obliged to spend at least 40 hours of continuing medical education (CME) per year. These CME-activities are provided by a number of organisations. On a national level, the Dutch College of General Practitioners (NHG) develops and distributes guidelines and educational materials based on these guidelines. These materials can be used for individual and for small-group education, and include varying methods like written materials or assessments. Distributed over the country, regional GP organisations are involved in CME and quality improvement. Activities differ strongly between regions, and include for example the organisation of courses, giving feedback on performance or support in specific areas of care, e.g. diabetes mellitus or COPD. Locally, GPs are usually involved in educational groups that focus on prescribing behaviour, practice organisation, peer review etcetera. These activities also differ strongly between local groups. Besides these professional organisations, private and commercial organisations can offer CME-activities. These are often initiated by the pharmaceutical industries.

The quality of CME courses is controlled by an independent organisation for accreditation and recertification of General Practitioners. GPs individually select CME activities out of the available supply on the basis of their own preferences or perceived learning needs. There are no demands regarding the content of the courses followed or the diversity of the programmes selected.

The study: general aim, research questions and design

The aims of the study were to develop an approach to support CPD and to investigate the effectiveness and feasibility of this approach in relation to the current routines. The project started in January 1999 and lasted until December 2003. More specific, the research questions were:

- What are the features of an approach that can be expected to support CPD of medical professionals effectively and efficiently?
- Is this approach more effective in changing GPs performances towards desired standards than current educational approaches?
- Is this approach feasible, i.e. what are the needs in terms of time and money and do GPs appreciate the developed approach?

The first step was the development of an educational approach. This was done on the basis of current evidence and assumptions in CPD. Also, the barriers encountered by

GPs in CPD in the current situation, and their preferences for future education were investigated. After outlining a global model for an educational intervention based on literature and on preferences of the target group, the specific requirements for this intervention were developed. These included a criterion-referenced standard for doctor-patient communication, software for giving feedback and training GPs to play specific roles within the intervention.

The educational intervention was subsequently investigated in the second part of the study period. To determine effectiveness, the design of a randomised trial was chosen. Learning activities in the broad domain of general practice were expected to lead to varying activities, which may result in improvements that remain hidden in global assessments of performance. To enable us to measure the effects, a subdomain of general practice was chosen. This subdomain existed of three topics: doctor-patient communication, ENT-disorders and type 2 diabetes. These represent both communicational and medical aspects of general practice. Furthermore, diabetes represents a chronic disease, while ENT-disorders represent more acute diseases. Participants in the intervention arm were offered the newly developed intervention, while participants in the control arm were offered existing educational materials on the three topics. Besides these activities, participants were free to chose CME-activities not related to the study. The effectiveness was measured on the level of performance, i.e. by focusing on GPs professional behaviour during daily surgeries. This was done using video-observation, data on prescriptions, prospective self-recording and data on practice organisation. The feasibility, in terms of the appreciation and of the needs to perform the intervention, was also investigated by time registration and by a survey.

Outline of the thesis

The following three chapters deal with the development of the educational intervention for continuing professional development of GPs.

Chapter 2 focuses on the concept of CPD and on the underlying theories and evidence. A format for an educational approach that supports GPs in their CPD is proposed. This format is based on available insights and evidence on education of practising physicians. Decisions on details in the intervention were made on the basis of preferences of GPs (see the study reported in chapter 3). A small-scale pilot study on the appreciation and feasibility of the intervention is described.

Chapter 3 describes a study towards the use and appreciation of methods for learning and improving by general practitioners. This study was needed to tailor the educational intervention to the needs and preferences of the target group.

Chapter 4 focuses on standard setting procedures for an assessment of performance in daily practice. In chapter 2 it is concluded that assessments are the starting point of the educational intervention, that assessments are preferably on the level of performance with the availability of a criterion-referenced standard, reflecting the desired performance of GPs. However, literature on standard setting of performance assessment is not available, nor is it known if and how a credible standard can be obtained.

The next chapters focus on the multifaceted approach that has become available. The effectiveness and feasibility were investigated, and are described in chapters 5, 6 and 7. In chapter 5, the effects on doctor-patient communication are depicted. Chapter 6 specifies the effects on ENT-disorders. The effects on the diagnosis and treatment of diabetes are not described, as only a small number of doctors were actively involved in learning activities, following the assessment results.

Chapter 7 describes aspects of feasibility, for example the costs of the intervention and the appreciation by the participants.

In the discussion of this thesis, the main results are summarised. The results and the stronger and weaker sides of the study are discussed. Finally, a few suggestions are made for future educational development and research on CPD.

Research questions

Chapter 2:

What are the characteristics of an approach aimed to support CPD of medical professionals that can effectively and efficiently improve the quality of care?

Chapter 3:

Do GPs use the available assessment methods; are assessment results used for CPD; what are the barriers and needs encountered using assessments and assessment results for CPD?

Chapter 4:

What is the applicability of common standard-setting procedures in authentic assessment i.e. to set a criterion-referenced standard for doctor-patient communication in daily care?

Chapter 5:

Is the developed CPD approach effective in improving doctor-patient communication?

Chapter 6:

Is the developed CPD approach effective in improving care for ENT-disorders?

Chapter 7:

What are the characteristics of the newly developed CPD-approach with regard to the feasibility and appreciation by stakeholders?

LITERATURE

1. Knowles MS. *Self-directed learning: A guide for learners and teachers*. Cambridge: The Adult Education Company, 1975.
2. Prochaska JO, Redding CA, Evers KE. The transtheoretical model and stages of change. In: Glanz K, Lewis FM, Rimer BK, editors. *Health behavior and health education*. San Francisco: Jossey-Bass Publishers, 1997:60-84.
3. Bero LA, Grilli R, Grimshaw JM, Harvey E, Oxman AD, Thomson MA. Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. The Cochrane Effective Practice and Organization of Care Review Group. *BMJ* 1998;317:465-8.
4. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing Physician Performance. *JAMA* 1995;274:700-705.
5. Oxman AD, Thomson MA, Davis DA, Haynes RB. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *CMAJ* 1995;153:1423-1431.
6. Thomson O'Brien MA, Freemantle N, Oxman AD, Wolf F, Davis DA, Herrin J. Continuing education meetings and workshops: effects on professional practice and health care outcomes. *Cochrane Library* 2001:1-42.
7. Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health technology assessment Winchester, England* 2004;8:iii-iv, 1-72.
8. Grol R. Improving the Quality of Medical Care: building bridges among professional pride, payer profit and patient satisfaction. *JAMA* 2001;286:2578-85.
9. Grol R, Grimshaw J. From best evidence to best practice: effective implementation of change in patients' care. *Lancet* 2003;362:1225-30.
10. Grol R. Beliefs and evidence in changing clinical practice. *BMJ* 1997;315:418-421.

Chapter 2

**DIRECTED
SELF-LEARNING
FOR MEDICAL
PROFESSIONALS:
DEVELOPMENT OF
AN INTERVENTION
TO SUPPORT CPD**

ABSTRACT

The effectiveness of current educational interventions to improve the quality of care in general practice has been found to be disappointing. We describe the development of a multifaceted intervention to support continuing professional development of Dutch GPs that aims to be more effective than current approaches, by tailoring structured learning activities to objective individual learning needs. The development was based on literature, on needs and barriers as encountered by GPs and on the opinions of experts involved in educating medical professionals.

Starting point for the development are the concepts and the underlying theories as described in literature on CPD. Combining these theories with the evidence on the effectiveness of educational methods, we propose directed self-learning as a concept that depends on individual professional initiatives, but also supports GPs in specific aspects of CPD.

The first step in our model is a needs assessment to support GPs in identifying what aspects of performance need improvement. On the basis of the assessment results and predefined criterion-referenced standards of adequate performance, a subdomain for further improvement is selected from the assessed domain. GPs receive feedback on their results and the selected subdomain. Consequently, they receive small-group education, supervised by a tutor who is an expert on the topic of interest.

In a small-scale pilot, this approach appeared to be acceptable and feasible for participating GPs, but some adjustments were necessary to enable the intervention to be used on a larger scale. The main change is to support GPs in the process of quality improvement only, but to consider the contents and targets of their activities as their own professional responsibility.

INTRODUCTION

This chapter describes the development of an educational intervention to support the continuing professional development (CPD) of General Practitioners (GPs) in The Netherlands. The development of such an intervention is necessary because, as was stated in the introduction of this thesis, maintaining the competence to practice is a critical component of professionalism, but the effectiveness of current educational approaches and methods has often been found to be disappointing.¹⁻³

The intervention was developed on the basis of current literature, and on the opinions of the target group and experts, i.e. GPs and those involved in their continuing education in The Netherlands. In this chapter the subsequent steps in developing the intervention are described in four separate sections. The first section focuses on the definition of CPD. Then, the second section elaborates on theoretical approaches associated with CPD. The third section depicts the intervention's model to support CPD, as well as the methods and instruments used, taking into account the needs and barriers expected by both experts and the target group (based on a study that will be described in more detail in chapter 3). Finally, the last section relates the first experiences with the intervention and the lessons learnt in a small-scale pilot.

This chapter does not aim to give a complete overview of current literature on learning and improving of physicians, since a number of carefully conducted reviews and valuable textbooks on this topic have become available during the last years. It mainly describes the concepts, evidence and ideas underlying the intervention, and accounts for the choices made in our specific context of CPD for Dutch GPs.

Throughout this chapter the term 'theories' refers to conceptual approaches towards learning of professionals (e.g. adult learning theory, problem based learning), the word 'methods' refers to specific types of educational activities (e.g. written materials, small-group working) and the word 'interventions' refers to the actual educational activities using single or combined methods.

DEFINITION OF CPD

In developing our intervention we started with focusing on the definition and theories available in literature on continuing professional development (CPD). To obtain relevant literature, a search was done in Medline, using the terms 'continuing professional development' and 'family practice' and related terms. An additional and broader search was performed to identify reviews on continuing education and improving the quality of care, by a Medline search aiming at these terms. Also, reviews concerning these fields that are available in the Cochrane library were collected. The obtained literature was scanned for references to relevant literature, i.e. referring to CPD, or to the development or effectiveness of educational interventions aiming at practising physicians. The articles as well as the textbooks and 'grey literature' on CPD were checked for relevant references.

We first studied the literature, focusing on the underlying concepts of CPD. It appeared that there is no commonly agreed definition of what CPD actually is, nor can it be clearly distinguished from other educational activities.^{4,5} Varying descriptions, focusing on different aspects of CPD are given in literature (box 1).⁶

Box 1. CPD in literature

Focus on the process

"Continuing Professional Development is the process by which health professionals keep updated to meet the needs of patients, the health service, and their own professional development. It includes the continuous acquisition of new knowledge, skills, and attitudes to enable competent practice. (...) The term continuing professional development acknowledges not only the wide ranging competences needed to practise high quality medicine but also the multidisciplinary context of patient care."⁴

"The process of CPD embodies the concept of lifelong continuing, continuous medical education within the context of continuing professional development."⁷

Focus on theories

"Continuing Professional Development implies a more learner centred and self directed approach to learning (compared to CME)."⁶

Focus on desired outcome

"Continuing professional development for healthcare professionals is an important strategic instrument for improving health (...) a way of maintaining standards of care; improving the health of the nation; and recruiting, motivating, and retaining high quality staff."⁸

CPD is "a process of lifelong learning for all, to meet the needs of patients and deliver NHS priorities".⁹

Focus on the content

"All (...) doctors need skills that extend beyond updating their medical knowledge in order to practise effectively (...). Such skills include management, education and training, information technology, audit, communication, and team building. These broader skills are embraced by continuing professional development (...)."⁵

"Continuing Professional Development is concerned with the acquisition, enhancement and maintenance of knowledge, skills and attitudes by professional practitioners."¹⁰

"CPD is a constant process of updating old skills and knowledge and gaining new ones."¹¹

"CPD is the post-registration acquisition of skills or knowledge by health-care professionals."⁸

The lack of consistency in the definitions of CPD may be a barrier for further development and research, by hindering discussion among educationalists or oversimplifica-

tion.¹² Therefore, we define CPD by integrating elements of the available definitions (box 2).

Box 2. Definition: Continuing Professional Development (CPD) is a doctor's ongoing learning process in order to meet medical and societal standards of desired performance in patient care.

In CPD, *development* is only *medical* (as in continuing medical education), but *professional*, which means, "belonging to or connected with a profession, which is a vocation or calling, especially one that involves some branch of advanced learning or science".¹³ The definition comprises the learning process of the individual doctor, and the aim of learning activities. This aim is related to medical and societal standards, i.e. it is based on the available scientific evidence but it is also linked to the context in which the professional works e.g. with regard to attitudinal aspects, financial restrictions or patients' demands. The word '*performance*' emphasises that educational activities should be directed to actual work with patients and colleagues in daily practice.¹⁴ Focusing on knowledge or competence, as some propose,¹¹ may distract the attention from performance, even though the former have been proven to be different concepts.^{15 16}

Conclusion

There is no commonly agreed definition of CPD in literature. We define CPD as the ongoing learning process of doctors to meet medical and societal standards in patient care.

THEORIES ON THE SUPPORT OF CPD

Using this definition as a starting point, we now focus on developing an intervention to support CPD, first of all studying theories on learning of professionals and on improving professional behaviour.¹⁷⁻²⁵ In this chapter, we restrict ourselves to the theories related to CPD; we refer to the textbooks for an overview of other theoretical approaches in the broader context of quality improvement in health care.^{17 18}

Theoretical approaches associated with CPD-support are the "learner-centred approach" and "self-directed learning".^{6 10 11 26} These terms are often used together, while a clear distinction between the two is seldom made. We discuss these approaches separately in more detail, considering them in the light of the evidence on effectiveness of educational activities that has become available during the last decades.^{1-3 27-29}

In order to facilitate the discussion of theoretical approaches to CPD-support, we use a metaphor that is consistent with our definition.

A journey towards ideal care

The metaphor of a journey to continuing professional development, made by every doctor during his or her career, has been used before in educational literature.³⁰⁻³²

The doctor is the traveller who departs from his or her actual practice and whose destination is "ideal care", in which all medical and societal standards are met. The aim of a CPD-intervention is to effectively and efficiently support the doctor in approaching this destination of "ideal care" as close as possible.

Learner centred approach

The learner centred approach originates from Problem Based Learning (PBL).³³ PBL intends to stimulate life-long, self-motivated learning, mostly in tutorials. The learning process focuses on the specific needs of individual learners and is adapted to their learning styles.²⁶ The learner and the specific context thus define contents and format of the learning activities.

A review on PBL in continuing education did not find convincing evidence for its effectiveness in experienced doctors, mainly due to the limited number of available studies.³⁴ More studies are available in undergraduate education, but here too, its effectiveness has been questioned.³⁵⁻³⁸ Given the methodological problems that arise in proving PBL's effectiveness, others claim that the focus should be on a further development and research of the assumptions underlying PBL.³⁹ In this view, PBL appears to be a rational approach in the context of current evidence, as it has several characteristics that are found to be related to effectiveness.

One of PBL's characteristics is the learner's active involvement in the educational process. Available evidence in the context of CPD shows that approaches that actively involve learners in the educational process, like small-group education or outreach visits, are more effective than teacher-centred methods, such as large scale lectures.^{1 2 28 29}

Another characteristic, tailoring learning to professionals' needs and contexts, also appears logical. Because of differences in professionals' competencies and in health care settings, the learning needs may vary. Even if the encountered problems are similar, differences in learning styles or barriers towards "ideal care" may require varying routes.¹⁸ Reviews have suggested that tailoring interventions has a positive influence on their effectiveness, but again, evidence is not straightforward.¹ Therefore, further development of PBL-approaches to support CPD and research on their effectiveness is necessary.

PBL-approaches may have disadvantages in terms of feasibility.^{38 39} Individualisation of learning paths may be effective, but educational interventions must also fit in the available time of learners and in the budgets and staff required. As both effectiveness and feasibility of approaches determine successful implementation, aspects of feasibility must be included in development and research of PBL for CPD.

In the view of Problem-Based Learning approaches, doctors have the same destination, but they depart from different places and encounter different barriers on

their journeys. Therefore they need individualised trajectories in order to approximate their destination of "ideal care".

Self-directed learning approach

"Self-directed learning (SDL) is the preparedness of a learner to engage in learning activities defined by himself rather than by a teacher. 'Preparedness' must be understood as having both a motivational aspect and involving skilled behaviour. Thus, an accomplished self-directed learner experiences an intrinsic need to acquire knowledge, not dominated by requirements set by his teacher. In addition, he has mastered the appropriate information seeking skills, that is: he knows where and how to find information resources that would fulfil his need."⁴⁰

SDL originates from adult learning theory, which has been balanced and criticised.^{34 41-}

⁴³ It can be doubted if doctors have the abilities needed for SDL, like assessing their learning needs and selecting appropriate learning materials and methods.

First, doctors have difficulties in identifying their learning needs.^{30 44} Self-assessment is required to consider stronger and weaker points in actual practice (metaphorically speaking: 'to determine their localisation'), but research in this field has shown disappointing results.⁴⁵⁻⁴⁷ 'To determine their destination', health care professionals need to be aware of what ideal practice looks like, while people tend to be unaware of what they do not know.⁴⁵⁻⁴⁹ Furthermore, doctors choose courses on the basis of many reasons other than objective or subjective needs.⁵⁰⁻⁵² Given the experience that doctors tend to follow courses on topics they are already interested in, self-directed learning may promote variation among doctors,^{23 45 53} contrary to the aim of CPD. This may be particularly problematic in professions that have to maintain numerous aspects of a broad task like general practitioners.⁵⁴

It must also be questioned if doctors can select educational methods for specific purposes appropriately.⁴² They have never been trained in doing this, and they have limited experience in educational techniques, as they are mostly educated in traditional curricula. It has been demonstrated that self-directed learning activities vary between professionals, depending on the underlying clinical problems, but the effectiveness of this selection has not been studied.^{55 56}

The concept of SDL also has characteristics that fit well in the specific context of CPD. First, it must be recognised that professionalism itself implies an important role and responsibility of professionals in taking care of their education. Also, SDL has advantages in terms of feasibility, as a lot of educational work is left to learners themselves instead of educational organisations, often struggling with a lack of time and money.

Self-directed learning implies that travelling doctors are able to determine where they are, where to go, how they should get there and which means of transport they should use. The initiative to travel in a specific direction lies with the doctors themselves, and is based on their own motivation.

Conclusion

In literature on CPD, two theoretical approaches are advocated: the learner-centred and the self-directed approach.

Characteristics of a learner centred approach that are expected to attribute to its effectiveness are: tailoring the intervention to learners' specific needs, and active involvement of learners in the learning process e.g. in small-group education.

A self-directed approach fits in the context of CPD because professionalism implies that learners have important responsibilities in their own learning processes, but there are serious concerns about the validity of self-assessment and the effectiveness of improvement strategies being self selected. However, with regard to feasibility self-directed approaches may have advantages as compared to problem-based approaches.

We hypothesise that an effective CPD-intervention should be based on a combination of aspects of both these theoretical approaches (PBL and SDL). A balance should be found between the professional responsibilities of learners in their educational processes on the one hand and the problems that health care professionals encounter in CPD on the other hand. Learners should therefore be supported in identifying learning needs and in using appropriate and effective educational methods. This may be characterised as "directed self-learning".⁵⁷

To ensure reaching the final destination, a doctor must have information on the questions: where am I, and what is my destination (ideal practice) and how do I reach this? Besides a description of his location and his destination to enable him to undertake the journey towards "ideal care", the travelling doctor needs additional support, e.g. means of transport, and a road map, or a timetable.

MODEL FOR A 'DSL-INTERVENTION'

In this section the concept of 'directed self-learning' is translated into a practical intervention to support general practitioners' CPD in The Netherlands. We designed an intervention containing three steps: assessment, feedback on assessment results and small-group meetings. In box 3, the educational support given to the GPs and their own responsibilities are briefly presented; these aspects will be gone into in more detail in the rest of this section.

Box 3.	Educational support	Professional responsibility
Assessment	<ul style="list-style-type: none"> - restriction of the domain of general practice to subdomains - definition of standards of desired care in these subdomains - identification of learning needs by relating assessment-results to standards 	<ul style="list-style-type: none"> - demonstration of professional competencies and current performances in patient care
Feedback	<ul style="list-style-type: none"> - revealing selected topics for improvement - motivating GPs for improvement 	<ul style="list-style-type: none"> - reflection on actual and required competencies, and current and desired performance
Small-group meetings	<ul style="list-style-type: none"> - context (colleagues and tutor) that may be motivating - content specific expertise of colleagues or tutor; shared activities in acquiring expertise - method (structured quality improvement) aiming at improving daily practice 	<ul style="list-style-type: none"> - definition of targets for improvement - initiating appropriate educational activities toward defined targets - improving actual patient care

Assessment

The first step in the intervention is to identify learning needs by comparing actual care with standards of desired care.^{30 26 45} GPs need support in gathering information on areas in which performance does not meet the standards.^{45 46 58} This can be given by objective measurements of performance, known as audit, appraisal, monitoring or assessment of actual practice.^{59 60} In this thesis we use the term 'assessment'.

For our intervention, we collected the available assessment methods for GPs in The Netherlands and selected the knowledge test and video-assessment of communication in GPs' own practices.⁶¹⁻⁷⁰ These were chosen considering the demands resulting from what was described above. These demands are: 'ideal care' is described, e.g. in a guideline; a standard for desired performance is defined and ideally assessment is on the level of performance, as CPD aims at actually delivered care. In the context of our study and for practical reasons, we decided to restrict the assessed domain to two to three topics, and to assess both medical and interpersonal aspects out of the broad domain of general practice.

Of the assessments available, no method met with all demands. A standard for desired performance for the video-assessment of doctor-patient communication was not available, and we therefore developed and studied criterion-referenced standards for this method (described in chapter 4). The knowledge test does not assess performance, but in previous studies, written knowledge tests have been preferred amongst assessments of lower levels of competence after a trade-off between predictive value and feasibility.^{15 69 71}

The identification of learning needs consists of two steps: (1) measuring performance and (2) deciding which aspects of care need improvement. We developed a procedure to select topics for improvement for each participant. Three aspects were involved in this procedure: assessment results, GP's motivation, and practical possibilities (used in this hierarchical order). The choice for improvement activities is primarily directed by the assessment results. In a number of cases however, no decision can be made, e.g. in a participant that scores below standards on all topics.⁷¹ In this case, the choice is made on the basis of GPs' motivation for the topic. Not until after the assessment, participants are given the opportunity to give a preference for a topic, as this may reveal aspects of desired care that the assessed GP was previously unaware of. If assessment results or motivation do not give a clear answer, the choice is made on the basis of practical possibilities.

We decided to make a choice on contents before participants received feedback. This was done on the basis of the findings from the study described in the next chapter. In the current context, assessment results are hardly used for targeted improvement. To stimulate the use of assessments, GPs suggested a more directed approach as a possible solution. Furthermore, assessments that are expensive in terms of time and money for both participants and supporting staff may in our view not be free of obligation.

Feedback

The second step in the intervention is to inform GPs on the assessment results and to motivate them for improvement; needs assessment alone is probably not sufficient to improve care.^{1 2 27 28 44 72-74} One possible reason is that GPs have difficulties with feedback on assessment results.^{69 75} Assessment must therefore be followed by clear and stimulating feedback.

In our intervention, assessment results and the selected topics for improvement were presented by a trained GP-colleague, who visited participating GPs in their own practices. Assessment results were presented visually as bar charts, comparing participants' results with the results of GPs in their own regions and with the standards of desired care. To identify specific topics for change, more detailed scores and standards on specific aspects of the assessed domains were given. Participants' acceptance of the assessment results and the motivation for improvement of the selected topics were discussed. Finally, advice was given on preparing for the first small-group meeting. This format was chosen on the basis of literature on feedback, and on a study of our target group's preferences.⁷⁵⁻⁸⁰ In our study, we found that GPs preferred written feed-

back, limited to one, or just a few pages, combined with an oral explanation. In literature, the key features of feedback are 'how', 'by whom' and 'when' it is given. Therefore, GPs were visited in their own practices to provide a 'safe' atmosphere. A GP-colleague as feedback giver enabled participants to discuss the information given. Assessment, feedback and consequent learning activities were organised in a limited time frame to enhance effectiveness.^{77 80}

Small-group meetings

The third step in the intervention is to support the translation of assessment results into effective improvements in care provision by a number of small-group meetings. We added this educational method on the basis of the needs reported by GPs,⁷⁵ and because we expected it to enhance the effectiveness of the intervention. By combining assessments with small-group meetings, a multifaceted intervention became available. Reviews available, suggested these to be more effective than single interventions,^{1 2 81} which has however not been confirmed by a more recent review.³

A number of possible methods are available to support the process of improving performance. Between these various methods differences in effectiveness have been found, yet considerable variation in effectiveness is also present within groups of similar methods.^{1 2 28 29 34 82 83} In the specific context of CPD, it has been stated that there is "no best learning method and no best approach to learning", and that preferably "doctors actual preferred ways of learning" are used.¹¹ Like other authors, we question if GPs can select educational methods for specific purposes appropriately, and therefore it was us who chose the format of small-scale meetings.⁴²

We chose small-group meetings for a number of reasons. First, they fit in the theoretical concepts previously discussed in this chapter. Also, small-group meetings can easily be tailored to the specific needs of a learner.¹⁷ Small-group meetings are effective, compared to other methods (like educational materials or large-scale educational meetings).^{1-3 29 82-85} Finally, small-group meetings are preferred by the members of our target group. These reported that they preferred learning and improving activities after assessments within small groups of GPs, that all focused on the same topic.⁷⁵

In the pilot, described in the following section, three small-group meetings were planned for each selected topic. The groups consisted of regional GPs and a local expert in the topic of interest to support learning.⁸⁶ The programme of the meetings needs to be decided upon by the participants and the expert themselves, as the specific improvement targets and the contents of the meetings are not known prior to the assessments.

Conclusion

The intervention that supports GPs in CPD in a way that can be characterised as 'directed self-learning', consists of a comprehensive format of assessment, feedback and small-group working. Assessment supports GPs in identifying aspects of performance in need of improvement. On the basis of the assessment results, topics for improvement are selected using a procedure that is primarily based on defined standards of desired

care. Feedback is given by a trained GP-colleague who visits participants in their own practices. Small-group meetings support GPs in improving performance in patient care. It is not clear if this intervention provides an appropriate balance between support and professional responsibility. The demands on the participants in terms of motivation and educational capabilities appear to be strong. Motivation, a key factor in theories on adult learning, may be limited in topics that are selected by assessments, but regular meetings in small groups may motivate learning and improving activities.³² It is not clear whether GPs are capable of shaping meetings on the basis of assessment results. Therefore, we investigated the feasibility on a small scale before studying the effectiveness of the intervention.

PILOT OF THE CPD-INTERVENTION

Introduction

A pilot study was performed to investigate questions that arose during the development of our intervention, described in the previous sections. These questions concerned the acceptance and the appreciation of the intervention by GPs. Furthermore, we needed to explore the feasibility of the approach before investigating its effectiveness on a larger scale.⁸⁷

The specific research questions were: do participants accept the criterion-referenced standards used; do participants accept the choices for topics to be made on the basis of the assessment results, instead of choosing themselves; are participants motivated and able to bear responsibility for the small-group meetings; what problems in feasibility are encountered in organising assessments, feedback and small-group meetings; what are the costs of the intervention for each participant?

Methods

The participants were eight voluntary GP's, out of two general practice groups ("HAGRO's"). The domain of the intervention was restricted to doctor-patient communication, and the diagnosis and management of asthma and COPD.

In the first step of the intervention, GPs were assessed on asthma and COPD, using a knowledge test, as well as on doctor-patient communication by means of video observation in daily practice. Every GP was allocated to one small group focusing on either of the topics, using the procedure described in the previous section. Each group had three sessions planned, which were guided by a tutor. In these meetings, a goal for improvement was defined; a plan for improvement of actual care was made and subsequently carried out.

The support-staff needed, consisted of one secretary to organise the assessments, two technicians to install the video-observation equipment, four GP-observers to score the video-taped consultations, one research assistant to prepare the written feedback, one GP giving feedback, one GP for the meetings on asthma/COPD and one psychologist for the meetings on doctor-patient communication.

An evaluative questionnaire was designed to investigate the appreciation and acceptance of the participating GPs. Aspects of feasibility in organising the intervention were evaluated by the support staff involved in the intervention. Costs of the intervention, as an aspect of feasibility, were determined as real costs of fees for all support staff involved.

Results

First, all GPs were assessed. Then, they were divided into 2 groups of 4 GPs on the basis of the assessment results. Assessment required a period of 4 weeks. Feedback was given in the week following the assessments, according to schedule. Improvement sessions started in the week following the feedback. The participating GPs attended all sessions.

Acceptance. Participants accepted both the criterion-referenced standards as well as the procedure to allocate GPs to small groups on the basis of their assessment results. Participants were capable and motivated to bear responsibility for the contents of the meetings: there were no difficulties in the contents of the improvement sessions, except for the first meeting of the group focusing on asthma and COPD. Formulating a specific goal for improvement turned out to be difficult for GPs in this group, which was attributed to the theoretical results of knowledge testing. An additional assessment of the quality of care in the treatment of asthma patients by self-registration was performed which provided material for the following sessions. All 8 GPs reported changes in their daily practice as a result of the meetings and 5 GPs regarded this method as more effective than usual CME.

Feasibility. In the trajectory starting with assessment until the first small-group meeting, all 8 participants and all 11 members of the support staff are involved. Their activities are interdependent, as small-group meeting can only start after completing assessments and giving feedback to all participants. We aimed to minimise the time between assessment and the first small-group meeting, to enhance effectiveness. This requires detailed planning, sufficient support staff (e.g. GP-observers for the videotaped consultations), and full cooperation of those involved in order to avoid big time gaps for the participants. Tutors and GPs reported they needed more support in the format of the meetings. As a possible solution, they suggested a more detailed programme of the specific goals of each meeting.

The costs were € 450,- per GP, mainly attributable to the video-observation.

Conclusions

An intervention, in which assessment, feedback and small-group education are combined, is a feasible and acceptable instrument in general practice. Criterion-referenced standards as a procedure to select topics for improvement on the basis of assessment results are accepted by GPs. GPs are motivated and capable to bear responsibility for the content of the small-group meetings. A short period of time between assessment, feedback and learning activities probably enhances effectiveness, but leaves little time

for supporting staff to prepare on the content. However, if they focus on the process of learning and improving, while the content of the meetings or visits is the responsibility of the professional involved, this problem can be solved.⁸⁸ For future use, a format of small-group meetings was chosen, guided by a GP-tutor instead of an expert on the specific content of the meetings. The GP-tutor is trained in supervising the group and in advising the participants on the process of stepwise quality improvement.

On the basis of the experiences in the pilot study, specific materials were developed. A strict schedule for the planning of all tasks was made, to restrict the time between assessment and improvement activities. To limit the time between assessment and feedback, software was developed that enables easy conversion of test results in written feedback. For both participants and tutors manuals were written which provided them with general information, diagrams to fill in and recognisable examples out of daily care, in order to support the format of the meetings. To define the content of their improvement trajectory, it is essential for materials that relate to the content of the assessed topics, such as written materials or supplementary assessment, to be available.

LITERATURE

1. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing Physician Performance. *JAMA* 1995;274:700-705.
2. Oxman AD, Thomson MA, Davis DA, Haynes RB. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *CMAJ* 1995;153:1423-1431.
3. Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health technology assessment Winchester, England* 2004;8:iii-iv, 1-72.
4. Peck C, McCall M, McLaren B, Rotem T. Continuing medical education and continuing professional development: international comparisons. *BMJ* 2000;320:432-5.
5. du Boulay C. From CME to CPD: getting better at getting better? *BMJ* 2000;320:393-4.
6. Cantillon P, Jones R. Does continuing medical education in general practice make a difference? *BMJ* 1999;318:1276-9.
7. Wentz DK, Paulos G. Is now the time for continuing medical education to become continuing physician professional development. *The journal of continuing education in the health professions* 2000;20:181-7.
8. Brown CA, Belfield CR, Field SJ. Cost effectiveness of continuing professional development in health care: a critical review of the evidence. *BMJ* 2002;324:652-5.
9. A review of continuing professional Development in general practice. London: Department of Health, 1998.
10. Grant J, Stanton F. The effectiveness of continuing professional development. A report for the chief medical officer's review of continuing professional development in practice.: The Joint Centre for Education in Medicine, 1998.
11. Grant J. *The Good CPD Guide*. Sutton: Reed Bussiness Information, 1999.
12. Campbell JK, Johnson C. Trend spotting: fashions in medical education. *BMJ* 1999;318:1272-5.
13. The Concise Oxford Dictionary.
14. Miller GE. The assessment of clinical skills/competence/performance. *Acad Med* 1990;65:S63-7.
15. Ram P, Van der Vleuten C, Rethans JJ, Grol R, Aretz K. Assessment of practicing family physicians: comparison of observation in a multiple-station examination using standardized patients with observation of consultations in daily practice. *Acad Med* 1999;74:62-69.
16. Rethans JJ, Sturmans F, Drop R, van der Vleuten C, Hobus P. Does competence of general practitioners predict their performance? Comparison between examination setting and actual practice. *BMJ* 1991;303:1377-80.

17. Grol R, Wensing M. *Implementatie. Effectieve verandering in de patientenzorg*. 2 ed. Maarssen: Elsevier gezondheidszorg, 2001.
18. Thorsen T, Mäkelä M. *Changing Professional Practice. Theory and Practice of Clinical Guidelines Implementation*. Copenhagen: Danish Institute for Health Services Research and Development, 1999.
19. Knowles MS. *Self-directed learning. A guide for learners and teachers*. Cambridge: The Adult Education Company, 1975.
20. Green L, Kreuter M, Deeds S, Partridge K. *Health education planning: a diagnostic approach*. California: Palo Alto: Mayfield Press, 1980.
21. Rogers EM. *Diffusion of innovations*. 3rd ed. New York: The Free Press, 1983.
22. Nowlen PM. *Approach to continuing education for business and the professions*. New York: MacMillan, 1988.
23. Fox RD, Mazmanian PE, Putnam RW. *Changing and learning in the lives of physicians*. New York: Praeger, 1989.
24. Fox RD, Bennett NL. Learning and change: implications for continuing medical education. *BMJ* 1998;316:466-8.
25. Bandura AB. *Social Foundations of thought and action: a social cognitive theory*. NJ: Englewood Cliffs N.J.: Prentice-Hall, 1986.
26. Bennett NL, Davis DA, Easterling WE, Friedmann P, Green JS, Koeppen BM, et al. Continuing medical education: a new vision of the professional development of physicians. *Acad Med* 2000;75:1167-72.
27. Beaudry JS. The effectiveness of continuing medical education: a quantitative synthesis. *The Journal of Continuing Education in the Health Professions* 1989;9:285-307.
28. Davis D, O'Brien MA, Freemantle N, Wolf FM, Mazmanian P, Taylor Vaisey A. Impact of formal continuing medical education: do conferences, workshops, rounds, and other traditional continuing education activities change physician behavior or health care outcomes? *JAMA* 1999;282:867-74.
29. Thomson O'Brien MA, Freemantle N, Oxman AD, Wolf F, Davis DA, Herrin J. Continuing education meetings and workshops: effects on professional practice and health care outcomes. *Cochrane Library* 2001;1:42.
30. Handfield-Jones RS, Mann KV, Challis ME, Hobma SO, Klass DJ, McManus IC, et al. Linking assessment to learning: a new route to quality assurance in medical practice. *Med Educ* 2002;36:949-58.
31. Wilkinson TJ, Challis ME, Hobma S, Newble DI, Parboosingh IJ, Sibbald RG, et al. The use of portfolio's for assessment of the competence and performance of doctors in practice. *Med Educ* 2002;36:918-24.
32. Bateman H, Kinmonth AL. Journeys and pathways: exploring the role of professional development advice and educational guidance for practitioners expressing interest in research. *Med Educ* 2001;35:49-55.
33. Schmidt HG. Foundations of problem-based learning: some explanatory notes. *Med Educ* 1993;27:422-32.
34. Smits PBA, Verbeek JHAM, de Buissonjé CD. Problem based learning in continuing medical education: a review of controlled evaluation studies. *BMJ* 2002;324:153-6.
35. Norman GR, Schmidt HG. The psychological basis of problem-based learning: a review of the evidence. *Acad Med* 1992;67:557-65.
36. Albanese MA, Mitchell S. Problem-based learning: a review of literature on its outcomes and implementation issues. *Acad Med* 1993;68:52-81.
37. Vernon DTA, Blake RL. Does problem-based learning work? A meta-analysis of evaluative research. *Acad Med* 1993;68:550-63.
38. Colliver JA. Effectiveness of problem-based learning curricula: research and theory. *Acad Med* 2000;75:259-66.
39. Norman GR, Schmidt HG. Effectiveness of problem-based learning curricula: theory, practice and paper darts. *Med Educ* 2000;34:721-8.
40. Schmidt HG. Assumptions underlying self-directed learning may be false. *Med Educ* 2000;34:243-5.
41. Norman GR. The adult learner: a mythical species. *Acad Med* 1999;74:886-9.
42. Eraut M. Do continuing professional development models promote one-dimensional learning? *Med Educ* 2001;35:8-11.
43. Baveye P. The adult learner: a misinterpreted species? *Acad Med* 2000;75:217-8.
44. Grant J. Learning needs assessment: assessing the need. *BMJ* 2002;324:156-9.

45. Tracey JM, Arroll B, Richmond DE, Barham PM. The validity of general practitioners' self assessment of knowledge: cross sectional study. *BMJ* 1997;315:1426-8.
46. Gordon MJ. A review of the validity and accuracy of self-assessments in health professions training. *Acad Med* 1991;66:762-9.
47. Woollicroft JO, TenHaken J, Smith J, Calhoun JG. Medical Students' Clinical Self-assessments: comparisons with external measures of performance and the students' self-assessments of overall performance and effort. *Acad Med* 1993;68:285-94.
48. Tamburrino MB, Lynch DJ, Nagel R, Mangel M. Evaluating empathy in interviewing: comparing self-report with actual behavior. *Teaching and learning in medicine* 1993;5:217-20.
49. Kramer AWM, Zuithoff P, Jansen JJM, Tan LH, Grol R, Vleuten Cvd. Growth of self-perceived clinical competence in postgraduate training for general practice. *submitted*.
50. Murray TS, Campbell LM. Finance, not learning needs, makes general practitioners attend courses: a database survey. *BMJ* 1997;315:353.
51. Woodcock A, White P, Smith H, Coles C, Campion-Smith C, Stannard T. GP selection of postgraduate education courses has implications for colleagues: messages for course providers and for those writing practice professional development plans. *Br J Gen Pract* 2000;50:785-90.
52. Pitts J, Vincent S. General practitioners' reasons for not attending a higher professional education course. *Br J Gen Pract* 1994;44:271-3.
53. Fox RD, Harvill LM. Self-assessments of need, relevance and motivation to learn as indicators of participation in continuing medical education. *Med Educ* 1984;18:275-81.
54. Springer MP. *Basic job description for the general practitioner*. Utrecht: LHV, 1981.
55. Slotnick HB. How doctors learn: physicians' self-directed learning episodes. *Acad Med* 1999;74:1106-17.
56. McClaran J, Franco E, Snell L. Type of clinical problem is a determinant of physicians' self-selected learning methods in their practice settings. *The Journal of Continuing Education in the Health Professions* 1998;18:107-18.
57. Lewkonja R. Medical revalidation and education--directed self-learning. *Med Educ* 2001;35:426-7.
58. Ram P. unpublished data.
59. Dauphinee WD. Revalidation of doctors in Canada. *BMJ* 1999;319:1188-90.
60. Southgate L, Pringle M. Revalidation in the United Kingdom: general principles based on experience in general practice. *BMJ* 1999;319:1180-3.
61. Pollemans M. Kennistoetsing bij Huisartsen. Universiteit Maastricht, 1994.
62. Jansen K. Toetsing van Technische Vaardigheden van Huisartsen. Universiteit Maastricht, 1998.
63. Van den Hombergh P. Practice Visits. Katholieke Universiteit Nijmegen, 1998.
64. Schuwirth L. An approach to the assessment of medical problem solving: computerised case-based testing. Universiteit Maastricht, 1998.
65. Wensing M. Patients evaluate General Practice. Katholieke Universiteit Nijmegen, 1997.
66. Spies T, Mokkink H. Eindrapport "Toetsen aan Standaarden". Nijmegen/Utrecht: WOK/NHG, 1999.
67. Jansen M, Jacobs A, Van der Weijden T, Grol R, Overdijk M, Van der Horst F. Toetsing van voorlichting door huisartsen aan patiënten. Maastricht: WOK, 1997.
68. Rethans J-J. Does Competence predict Performance? Universiteit Maastricht, 1991.
69. Ram PM. Comprehensive Assessment of General Practitioners. Universiteit Maastricht, 1998.
70. Ram P, Grol R, Rethans J-J, Schouten B, Van der Vleuten C, Kester A. Assessment of general practitioners by video observation of communicative and medical performance in daily practice: issues of validity, reliability and feasibility. *Med Educ* 1999;33:447-454.
71. Ram P, Vleuten Cvd, Rethans J, Schouten B, Hobma S, Grol R. Assessment in general practice: the predictive value of written-knowledge tests and a multiple-station examination for actual medical performance in daily practice. *Med Educ* 1999;33:197-203.
72. Sibley JC, Sackett DL, Neufeld V, Gerrard B, Rudnick KV, Fraser W. A randomized trial of continuing medical education. *N Engl J Med* 1982;306:511-5.
73. Borgiel AE, Williams JI, Davis DA, Dunn EV, Hobbs N, Hutchison B, et al. Evaluating the effectiveness of 2 educational interventions in family practice. *CMAJ* 1999;161:965-70.
74. Norman GR, Shannon SI, Marrin ML. The need for needs assessment in continuing medical education. *BMJ* 2004;328:999-1001.

75. Hobma S, Ram P, van der Vleuten C, Grol R. Gebruiken huisartsen toetsen voor kwaliteitsverbetering? *Huisarts & Wetenschap* 2002;45:726-9.
76. Ende J. Feedback in clinical medical education. *JAMA* 1983;250:777-81.
77. Hewson MG, Little ML. Giving feedback in medical education: verification of recommended techniques. *J Gen Intern Med* 1998;13:111-6.
78. Maguire P, Roe P, Goldberg D, Jones S, Hyde C, T OD. The value of feedback in teaching interviewing skills to medical students. *Psychol Med* 1978;8:695-704.
79. Rolfe I, McPherson J. Formative assessment: how am I doing? *Lancet* 1995;345.
80. Brinko KT. The Practice of Giving Feedback to Improve Teaching. *Journal of Higher Education* 1993;64:574-592.
81. Wensing M, Grol R. Single and combined strategies for implementing changes in primary care: a literature review. *Int J Qual Health Care* 1994;6:115-132.
82. Thomson O'Brien MA, Oxman AD, Davis DA, Haynes RB, Freemantle N, Harvey EL. Audit and feedback versus alternative strategies: effects on professional practice and health care outcomes. *The Cochrane Library* 2001:1-12.
83. Thomson O'Brien MA, Oxman AD, Haynes RB, Davis DA, Freemantle N, Harvey EL. Local opinion leaders: effects on professional practice and health care outcomes. *The Cochrane Library* 2001:1-30.
84. Freemantle N, Harvey EL, Wolf F, Grimshaw JM, Grilli R, Bero LA. Printed educational materials: effects on professional practice and health care outcomes. *The Cochrane Library* 2001:1-24.
85. Grimshaw J, McAuley LM, Bero L, Grilli R, Oxman AD, Ramsay C, et al. Systematic reviews of the effectiveness of quality improvement strategies and programmes. *Qual Saf Health Care* 2003;12:298-303.
86. Field S. Continuing professional development in primary care. *Med Educ* 1998;32:564-6.
87. Hobma S, Ram P, van der Vleuten C, Grol R. Feasibility and acceptability of an integrated system of assessment and improvement on asthma/COPD and doctor-patient communication for general practitioners. *European Journal of General Practice* 2000;6:A.18.
88. Beyer M, Gerlach FM, Flies U, Grol R, Król Z, Munck A, et al. The development of quality circles/peer review groups as a method of quality improvement in Europe. *Fam Pract* 2003;20:443-51.

Chapter 3

GPS' USE OF ASSESSMENTS FOR CONTINUING PROFESSIONAL DEVELOPMENT

ABSTRACT

Background

Assessment is a key element of continuing professional development, but little is known about its current use and educational impact. We studied whether Dutch general practitioners use the available assessment methods, if they use feedback for CPD, and what barriers they encounter.

Methods

The actual use of assessment methods in daily practice and the barriers encountered in their use were investigated by interviewing a randomly selected group of general practitioners by telephone.

The use of feedback on assessment results for CPD was investigated with a written questionnaire, sent to general practitioners with extensive experience with assessments.

Results

Of the 218 general practitioners asked to participate, 106 (49%) co-operated. More than half of them had little or no experience with assessment. Barriers were lack of time, problems in the cooperation with colleagues, difficulties in making a start, lack of motivation and self-discipline and lack of legislation.

The written questionnaire, sent to 90 general practitioners, was filled in by 82 GPs (91%). Eighty-two percent of this group had not used the feedback for CPD. Barriers were lack of time, problems in the practice organisation or cooperation with colleagues, interpreting the feedback, and lack of skills in CPD.

Conclusions

The use of assessment methods for CPD in Dutch general practitioners is limited. According to GPs, this is due to barriers in using assessment methods and barriers in using the results. Workload reduction, improvement of practice organisation, enhancing the support of GPs in using the methods, and legislation may increase the use of assessment for CPD.

INTRODUCTION

Continuing education for practising general practitioners (GPs) has shown great changes during the last decades, based on the growing insights in the effectiveness of educational activities.¹⁻³ The current approach to education for health care professionals, referred to as continuing professional development (CPD), is more learner centred and focuses on daily practice. To support the process to keep daily practice updated, a number of different instruments and methods, e.g. guidelines and courses, are available.

Usually, educational activities of GPs are self-directed, i.e. doctors select their own topics for continuing education. The effectiveness of self-directed activities and the required self-assessment have been seriously criticised.⁴⁻⁶ Objective assessment is necessary for taking purposeful action on the real weak spots.^{4,5} To improve the effectiveness of CPD-activities, the policy of educational developments have set out a central role for assessment since more than 10 years.

GPs' use and opinions about methods for CPD were examined in the early nineties.⁷ GPs had a positive attitude, but they had little experience with the then new instruments and methods. Since then, a wide range of assessments has become available.⁸⁻¹⁰ However, availability does not mean that the instruments are used neither that the feedback is an effective stimulus for CPD activities. Therefore, we performed this study in which we investigated the following questions. Do GPs use the currently available assessments? What is the attitude of users and non-users of available instruments? Is feedback used for quality improvement? What are the barriers, and how can the use of the assessments be further developed?

METHODS

To investigate the use of assessments for CPD, two studies were done. In the first study we investigated the use of the available instruments and methods and the barriers encountered in a random population. This was done with a semi-structured interview by telephone.

Because assessment can only be effective if the feedback is used for further CPD-activities, we did a second study focusing on GPs who had experience with almost all available assessments. They were asked to about their experiences in using the feedback for CPD by a written questionnaire.

Study 1. The use of assessments

The Netherlands Institute for Health Services Research (NIVEL) took a representative sample from the national register of established GPs. The GPs received written information about the contents of the interview. Trained interviewers phoned them and if they agreed to participate, an appointment for the interview was made.

The protocol for the interviews was based on data from the literature.^{1,3,7} The first part consisted of open questions, for example, 'With which assessments have you had ex-

perience?' and 'Did you experience difficulties in using them?'. Assessment methods that had not been named were explained and discussed. Barriers that the GPs had come up against or expected were inventoried. Then the question of possible solutions arose. Finally, a number of suggestions were propounded. In the last part of the interview, a number of closed questions were asked. The interview took 20 minutes. The answers were transcribed and categorised to match a previous investigation.⁷ The respondents were compared to the national population of GPs using the Chi-square test. The answers on the closed questions regarding attitudes and barriers encountered by frequent users of assessments and non-frequent users were compared using the Chi-square test.

Study 2. The use of feedback

A questionnaire to be filled in anonymously was sent to all the participants of a project, in which participants had experience with almost every available form of assessment, more than a year after it concluded.¹¹ The questionnaire consisted of closed, multiple-choice questions and was developed from data in the literature.¹²⁻¹⁵ The questions ask whether objectives for quality improvement have been chosen from the feedback, what activities have been undertaken, and what hindrances and requirements they experienced. The addressed GPs were compared to the national population of GPs using the Chi-square test.

RESULTS

Study 1. The use of assessments

Two hundred and eighteen GPs were approached, and 106 of them took part in the interview (response 49%). Reasons for refusing cooperation were: not interested (45%), no time (44%), or other reasons, such as absence, no remuneration, or ending the practice (10%).

A comparison of the representativeness of the sample of 106 GPs with the national population of GPs revealed no significant differences with respect to sex, mean age, urban area and single-handed working GPs vs. non-single handed GPs.

The number of assessments that GPs had experience with are given in table 1. Table 2 shows the methods named. Table 3 shows the barriers that GPs encountered or reported to expect in using assessments.

Table 1. The number of assessment methods used by a random group of GPs, interviewed by telephone (n=106).

Number of assessments used	Number of GPs
0	24 (23%)
1	35 (33%)
2	31 (29%)
3	10 (9%)
4	4 (4%)
5	1 (1%)
6	1 (1%)

Table 2. The assessment methods used by a random group of GPs, interviewed by telephone (n=106).

	Number of GPs
Intercollegial assessment	69
Written knowledge assessments	27
Assessment according to NHG-guidelines	15
Prescription data	8
Practice visits	7
Inspection visits to the practice	6
Checking patient data	5
Assessing knowledge via diskettes/internet	4
Assessments of knowledge on skills	3
Data of sick funds	3
Skills assessment	3
Diagnostic peer review	2
Patients' opinions	0

The most important barrier was lack of time due to the pressure of work. Solutions were especially aimed at reducing this. Also, more support for assessment, e.g. a regional facilitator, was seen as a solution.

The organisation and cooperation within the practice or the local general practice group was often seen as a hindrance to assessment. This concerned organisational limitations and differences in attitude and skills between colleagues. A better organisation of practice and of the local general practice group, more guidance by professional organisations, regulations, remuneration, and the creation of special assessment groups were named as solutions for these problems.

Table 3. The barriers that a random group of GPs, interviewed by telephone, experienced or expected in using the available assessment methods (n = 106).

	GPs (%)
<i>In the GP's surroundings and work situation</i>	
No time/too much work pressure	72
Practical problems with the organisation of the practice	22
Negative attitude from colleagues	21
Negative attitude from other care givers and specialists	3
Practical organisation problems within the local general practice group	2
Negative attitude from patients	1
<i>Related to assessments and the organisation of assessment</i>	
No faith in the existing assessments	14
Does not know how to begin	7
Not enough regulation/too little obligation	6
Lack of clear follow-up	2
Lack of effect in the practice	2
<i>Related to the individual GP</i>	
Difficulty changing old routines	9
Not motivated to change	4
Assessment is threatening	3
Age	2

A majority of those interviewed felt that the role of assessment should be stressed more, with the professional organisations in a leading role, and that there should be stricter regulations. Organising this should be in the hands of an independent organisation, for example, the Dutch College of General Practitioners. Some GPs would prefer to see a completely independent organisation.

More than half the GPs agreed with the suggestion that the compilation of the CPD-programme should be guided by the assessment results. Just as many GPs agreed with the suggestion that assessment results should be considered at re-registration.

It was not clear to one-third of the interviewed GPs just where the assessments come from. Just as many were concerned with what is done with the assessment results, and they were afraid that health care insurers would abuse the data.

Many of those interviewed did not know how to get started with assessments. More information and better accessibility to the methods were put forward as solutions. The interpretation of the feedback was also said to be a problem.

Some of the GPs were unwilling or did not have the skill to implement improvements that the assessment results led to. Some GPs had difficulty with being judged by colleagues. The creation of a safe debriefing with the assessment results and of assessment groups was proposed as a solution. These GPs often saw hindrances in the organisation

of the practice and the local general practice group and did not consider this a safe group. A third barrier was the lack of self-discipline. Some GPs said that their own age was a stumbling block. There were no differences in attitudes towards assessment or in the (expected) barriers between GPs who had little or no experience with assessment and GPs who had used several assessment methods.

Study 2. The use of feedback

The use of assessment results for CPD was investigated by means of a written questionnaire, which was sent to 90 GPs, 82 of who returned it (response 91%). The addressed GPs were representative for Dutch GPs with respect to sex, percentage of single-handed GPs and urban area. The study population was slightly younger than the national population (43.1 vs. 44.3 years; <0.05).

Eighty-two percent of respondents had not used the results for the choice of CPD-activities. Twelve percent used the feedback to draw up a personal course of study. The GPs would have liked to be able to do more with the feedback. Table 4 shows the hindrances found in using the feedback.

Again, lack of time was named as the main hindrance. Also, shortcomings in practise organisation and imperfect cooperation with colleagues were mentioned as a barrier. The nature of the feedback appeared to be an important reason for not using it. The feedback consisted of tables with written explanations, with details about hundreds of items. Eighty-two percent found it difficult to get an overview of the strong and weak points because of an information overload. Seventy-four percent saw a one-page summary as a solution.

Table 4. The barriers encountered by GPs (n=82) experienced with almost all available assessment methods, in using feedback on assessment results.

	GPs (%)
<i>In the surroundings and work situation of the GP</i>	
No time/too much work pressure	63
Practical problems with the organisation of the practice	53
Sub optimal cooperation with colleagues	10
<i>Related to assessments and the organisation of assessment</i>	
Unclear feedback	82
<i>Related to the individual GP</i>	
Do not know how to begin /no skills for systematic quality improvement	60

Sixty percent of the GPs were of the opinion that they did not possess enough of the required skills to take the step from assessment results to CPD. They were not sure how to tackle this problem. The solutions proposed, were that help and guidance should be offered after the assessment. Most (69%) had no clear preference for the form in which

this should take place. Thirty-one percent of the respondents preferred cooperation in a group of GPs, 25% wanted to do this with GPs who were working on the same subject, and 6% chose the local general practice group for this purpose.

DISCUSSION

The role of assessment in current CPD of GPs is disappointing. Many GPs hardly use the available methods. GPs who concern themselves actively with assessment only use the feedback in a limited way for CPD.

Caution is recommended for the interpretation of the data, which are not representative of Dutch GPs. The response to the interviews was less than 50%. The non-respondents often refused to participate because of 'lack of interest' and their opinions about assessment are possibly more negative than those of the respondents. A selected, motivated group was polled with a written survey. Taking these limitations into consideration, a large group of GPs have given data about requirements and hindrances, which is valuable for insight into the current state of affairs and for the policy on CPD.

It is remarkable that GPs see assessment positively, but use it sparingly. The perceived barriers are mainly of a practical nature, and they have changed little in comparison to those of a previous investigation.^{7 12} Lack of time is given as the most important hindrance. Further, the availability, the feedback, and the assistance in assessment are inadequate at the moment. GPs are also afraid that insurers can misuse assessment results. Only a limited number mention motivation and self-discipline as hindrances.

The suggested solutions require great changes in general practice and in the organisation of CPD. For example, the desired reduction of the pressure of work will probably not be achieved in the short term due to the threat of a shortage of GPs. It is also striking that few GPs see the local general practice group as a suitable group to process the assessment results, while health care policy wants these groups to play a more important role in CPD.

It is clear that developing assessments and making them available is not enough. At present, the assessments are an isolated instrument in the large range of educational courses offered. Rethinking the role and the form of assessment in CPD is urgently wanted. Before new policy and instruments are developed, it is necessary to investigate the form in which assessment and CPD must be brought about, and the feasibility and effectiveness of the new form.

LITERATURE

1. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing Physician Performance. *JAMA* 1995;274:700-705.
2. Davis D, O'Brien MA, Freemantle N, Wolf FM, Mazmanian P, Taylor Vaisey A. Impact of formal continuing medical education: do conferences, workshops, rounds, and other traditional continuing education activities change physician behavior or health care outcomes? *JAMA* 1999;282:867-74.
3. Oxman AD, Thomson MA, Davis DA, Haynes RB. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *CMAJ* 1995;153:1423-1431.
4. Gordon MJ. A review of the validity and accuracy of self-assessments in health professions training. *Acad Med* 1991;66:762-9.
5. Tracey JM, Arroll B, Richmond DE, Barham PM. The validity of general practitioners' self assessment of knowledge: cross sectional study. *BMJ* 1997;315:1426-8.
6. Norman GR. The adult learner: a mythical species. *Acad Med* 1999;74:886-9.
7. Grol R, Wensing M. Implementation of quality assurance and medical audit: general practitioners' perceived obstacles and requirements. *Br J Gen Pract* 1995;45:548-52.
8. Ram P, Grol R, Rethans J-J, Schouten B, Van der Vleuten C, Kester A. Assessment of general practitioners by video observation of communicative and medical performance in daily practice: issues of validity, reliability and feasibility. *Med Educ* 1999;33:447-454.
9. Wensing MJP, Van der Vleuten C, Grol R, Felling AJA. The reliability of patients' judgements of general practice care: How many questions and patients are needed? *Qual Health Care* 1997;6:80-85.
10. Van den Hombergh P, Grol RPTM, Van de Hoogen HJM, Van den Bosch WHM. Practice visits - Assessing and improving management in general practice. *Br J Gen Pract* 1998; 48:1743-1750.
11. Ram PM. Comprehensive Assessment of General Practitioners. Universiteit Maastricht, 1998.
12. Balas EA, Boren SA, Brown GD, Ewigman BG, Mitchell JA, Perkoff GT. Effect of physician profiling on utilization. Meta-analysis of randomized clinical trials. *J Gen Intern Med* 1996;11:584-90.
13. Berwick DM. Developing and testing changes in delivery of care. *Ann Intern Med* 1998;128:651-6.
14. Eliasson G, Berg L, Carlsson P, Lindstrom K, Bengtsson C. Facilitating quality improvement in primary health care by practice visiting. *Qual Health Care* 1998;7:48-54.
15. Wolfersteig J, Dunham S. Performance improvement: a multidimensional model. *Int J Qual Health Care* 1998;10:351-4.

 

Chapter 4

SETTING A STANDARD FOR PERFORMANCE ASSESSMENT OF DOCTOR-PATIENT COMMUNICATION IN GENERAL PRACTICE

ABSTRACT

Objective

Criterion-referenced standards for assessing performance in real practice of GPs should be available to identify learning needs or poor performers for CPD. The applicability of common standard-setting procedures in authentic assessment has not been investigated.

Methods

To set a standard for assessment of GP-patient communication with video observation of daily practice, we investigated two well-known examples of two different standard setting approaches. An Angoff procedure was applied to eight written cases. A borderline regression method was applied to videotaped consultations of 88 GPs. The procedures and outcomes were evaluated by the applicability of the procedure, the reliability of the standards and the credibility as perceived by the stakeholders, namely, the assessed GPs.

Results

Both methods are applicable and reliable. The Angoff standard for an adequate GP was 3.1 on the MAAS-Global scale ranging from 0 to 6. The borderline regression standard for an adequate GP was 2.5. The RMSE of the Angoff procedure was 0.12 on the MAAS-Global scale; the RMSE of the borderline regression procedure was 0.05. The obtained standards are credible according to the assessed GPs.

Conclusions

Both modified methods can be used to set a standard for assessment in daily practice. The context in which the standard will be used, e.g. the specific purpose of the standard, the moment the standard must be available or if specific feedback must be given, is important because methods differ in practical aspects.

INTRODUCTION

Practising doctors need continuing professional development (CPD) to sustain or improve the outcomes of their daily work. Assessment is needed to identify learning needs for CPD.^{1,2} This assessment needs to be as authentic as possible, and preferably relate to performance in daily practice, or in other words, to be at the "do" level at the top of Miller's pyramid.^{3,4} We need a minimum standard of desired performance to identify topics for improvement or individual professionals in need of remediation. Some standard-setting procedures, particularly in the area of written assessments, have been described in the assessment literature.⁵⁻⁷ More recently research has been focussed on setting standards for objective structured clinical examinations (OSCE's).⁸⁻¹¹ Setting standards in performance assessment has not yet been well studied.

A distinction is made between relative and absolute standards in the literature. A relative or norm-referenced standard is based on the scores of a group of test-takers, e.g. the 15% bottom group. An absolute or criterion-referenced standard is based on the mastery of the content of the assessment, that is, how many items or tasks have to be performed well to pass the examination. We prefer criterion-referenced standards in CPD because assessed groups may differ in ability for and in mastery of different topics.¹²

It is not yet clear whether existing standard setting methods can be adopted for performance assessment. Methods for obtaining criterion-referenced standards can be classed as test-centred and examinee-centred methods. In test-centred methods, expert judges rationally analyse the content of the test, whereas in examinee-centred methods, expert judgement is based on the candidates' performance. A popular method representing the former class is the *Angoff* method; examples of the latter class are the *borderline* and *contrasting groups* methods. These methods use 'borderline' performance as an anchor point for the minimum standard. For CPD purposes, we advocate a more stringent approach in which the minimum standard would preclude as many false-positive decisions as possible, since the purpose of CPD is to enhance performance to a level where adequacy is indisputable. Therefore, we consider borderline performance as 'to be improved'.

It is known that different methods lead to different standards, and a 'gold standard' of criterion performance is lacking. Therefore, the credibility of the methods has to be defined by other criteria related to the procedure and the outcome.¹³ These criteria depend on the purpose and the specific context of the assessment.

We are not aware of any previous studies investigating the credibility of standard-setting procedures for assessing performance in daily practice. Recently, several authors have emphasised the need for these standards.¹⁴ In this chapter, we study two criterion-referenced standard-setting methods: a modified *Angoff* method and a modified *borderline* method, both applied to an assessment of performance in the domain of general practice. These methods are used to set a standard on an important subdomain in general practice, namely, doctor-patient communication. For the exact procedure of assessing communication performance in daily practice, in which consultations are videotaped

during daily surgeries and scored on the MAAS-Global scale, we refer to an earlier publication.¹⁵ The study was undertaken to answer the question whether these standard-setting methods can be applied to set standards for performance assessment.

The credibility of the standard-setting procedures is studied using three criteria for credibility, relevant at least in our context of CPD. The first criterion is the *applicability* of procedures normally used in standardised test-taking conditions for the purpose of assessing performance in practice. In performance assessment, the content is not fixed, and the conditions under which the assessment takes place are not standardised. This might particularly be problematic for the content-based standard-setting methods. A second, rather generally accepted criterion for credibility is the *reliability* of the procedure. Credibility is also determined by the extent to which the standard is realistic.¹³ A criterion that is often used is the pass or fail rate. However, in identifying topics for CPD where performance might be very good or very poor, we cannot state any defensible pass or fail rates in advance. In CPD, it is vital that an assessment outcome is acceptable to its stakeholders for the assessment to be a stimulus for CPD activities. Therefore, we used the GPs' *perceived credibility* of the pass or fail decision after feedback on passing or failing as a third criterion.

These three criteria are used as main criteria for credibility. For practical reasons, we would prefer a standard that is known prior to the testing to give feedback shortly after assessment. We also prefer a standard on details of the assessment to give specific feedback that allows GPs to choose topics for improvement of their performance. Lastly the amount of time and money that has to be invested to obtain the standard is an important characteristic of the procedures.

METHODS

First, the assessment procedure for doctor-patient communication in general practice is explained. Then the two standard-setting procedures are described. The evaluation of credibility by the three criteria is discussed last.

Video assessment

A video observation system, consisting of two cameras with microphones, is installed in the GP's consultation and examination room for one week. All consultations for which patients give consent are recorded. A trained GP observer selects consultations and applies the MAAS-Global scoring system. This scoring list contains 3 sections: communication skills relevant to the specific phases of the consultation, general communication skills and medical aspects. The first two sections, containing 13 items on doctor-patient communication are the focus of the current study. The rating list, a summary of all items, used in scoring consultations, is given in the *appendix*. The items are case independent and global, but anchored with detailed criteria and scored on a seven-point scale.¹⁶ In giving feedback to a GP, the score per item is defined as the average of the GP's scores on the item across all cases. The score for a case is defined as the mean of

all items across one case. The score on a set of cases, the total score, is defined as the mean across the sample of cases.

The Angoff method

An Angoff procedure was adapted by applying a set of hypothetical written cases. The content of these eight cases, written by a panel of seven GPs, was based on a blueprint and reflected common complaints in general practice.^{4 17} The cases represent acute and chronic diseases, as well as single or multiple problems within the same case. They contain all relevant medical, psychological and social information from the patient's point of view in detail. Each case starts with a description of personal characteristics, the patient's family and medical history and current medication. Then the complaint is dealt with in more detail. Relevant aspects, such as information on work or school, the reason for encounter and the patient's expectations are described in the next sections. Finally, the information on a fictitious patient file is given. The cases are derived from role-playing scripts for simulated patients and their content has been validated in a previous study.⁴ The set of cases meets the criteria used for the selection of consultations in video assessment.

A panel of eight were invited to be judges. They were selected to fulfil the following criteria:

1. being a GP or senior trainee in general practice;
2. being interested in doctor-patient communication;
3. having experience in observing videotaped GP consultations, and
4. being familiar with MAAS-Global scoring.

The judges were 4 full-time working GPs, 2 part-time GPs and trainers in general practice, 1 part-time GP and member of the Dutch College of General Practitioners and 1 senior trainee in general practice.

The panel started with a discussion on quality of performance. A classification for practicing GPs in 5 categories (inadequate, marginal, adequate, good, outstanding) was used, analogue to accepted definitions of performance from the Dutch College of General Practitioners. Subsequently, the middle category of an 'adequate GP' – not a marginal GP – was chosen and defined as an anchor point for minimum performance, and was deemed appropriate for selection for CPD activities. Following this discussion, the judges gave their estimates. First judges were asked to read the case carefully and consequently estimate the probability of a satisfactory score (defined as 4 or more on the 7-point scale of the MAAS-global) of an 'adequate GP' meeting this patient during daily surgeries on each item of the scoring list. A reality check was introduced by allowing the judges to change their estimates as a result of discussing one another's judgements and by comparing the resulting minimum performance standard on real life data obtained in a previous study.¹⁵ This iterative procedure was done consecutively for all eight cases in a 7-hour meeting. The final minimum standard was the average of all estimates across items and cases.

The borderline regression method

Recently, an adaptation of the borderline method, which has advantages over the classic borderline approach, has been proposed.⁹ In addition to the MAAS-global scoring list, observers complete a rating scale ranging from inadequate to outstanding instead of giving a single borderline judgement. The best-fitting linear relation of MAAS-global scores and ratings is estimated. Using the obtained linear equation, the cut-off point on the score scale is estimated by substituting the a priori known cut-off point for adequate performance on the rating scale. We used a five-category classification (inadequate, marginal, adequate, good, outstanding), similar to the categories in the Angoff method, and we again chose the 'adequate GP' as an anchor point for the desired minimum cut-off point.

The judgements were based on 704 selected video-recorded consultations of 88 GPs observed during 1 week in daily surgeries. The study population was representative of Dutch GPs on sex; more GPs were working part-time (55% vs. 32%; $p < 0.05$), more were working in a group-practice (24% vs. 15%; $p < 0.05$) and less as solo GP (28% vs. 44%). The study population was younger than the Dutch GPs (mean age 44.4 vs. 47.6 years). The criteria used for the selection of 8 consultations for every GP were: (1) only initial consultations, (2) at least 6 cases with duration ≥ 7 minutes and ≤ 15 minutes/case, (3) no cases with more than 2 reasons for encounter, (4) age distribution: youngest patient < 18 and oldest patient > 65 years, (5) gender heterogeneity: at least 3 male or female patients, (6) representation of at least 6 different ICPC chapters, (7) representation of 5 ICPC chapters with highest prevalence (D, K, L, R, S), (8) the first 5 videotaped consultations were not to be included. A single observer rated each set of 8 consultations per GP. Consequently, one over-all judgement of the observed GP according to the five-scale classification (inadequate, marginal, adequate, good, outstanding) was made. In total, 41 GP judges were involved. They had been trained in two 3-hour sessions.

The standard was obtained in the following steps:

1. The global rating and average total score were obtained for all GP candidates.
2. A straight line was fitted by regressing the average total score to the global rating of the GP candidates.
3. The cut score was calculated from the total score corresponding to the rating of the level 'adequate' performance, according to the fitted straight line obtained in step 2.

Procedure

After their performance was assessed with video observation in their own practices in the procedure just described, two groups of GPs who were involved in a randomised controlled trial (RCT) on the effectiveness of a CPD approach were confronted with the Angoff standard or the borderline regression standard. The GPs, who participated voluntarily, were randomly assigned to either group. The Angoff standard was used in the intervention arm of the study, in a CPD programme consisting of assessment, personal

feedback and consecutive improvement. For practical reasons, the Angoff standard was appropriate for the experimental group in the RCT because it was available in advance. The borderline standard was given as written feedback after the assessment without any further planned CPD activities, as this was the control condition in the RCT. For GPs in both groups, feedback consisted of the individual total score, the pass or fail decision based on the respective standard for that group and the scores of the entire group as well as scores on separate items averaged across the cases.

Statistical analysis

The judges for both methods were instructed to make judgements only if they felt capable of doing so. *Applicability* was inferred from verbal comments of the judges that they felt comfortable about making these judgements. This issue was addressed during the meeting for the Angoff method, and it was done on an individual basis for the borderline regression procedure.

The *reliability* of the Angoff procedure was investigated on the basis of the root mean square error (RMSE) of the procedure.^{9 10} The RMSE is an estimate of the standard error of the cut score contained in the standard-setting procedure. The RMSE was estimated with a generalisability analysis (GENOVA), according to the following formula:

$$RMSE_{ANG} = \sqrt{\frac{\sigma_j^2}{n_j} + \frac{\sigma_c^2}{n_c} + \frac{\sigma_{jc}^2}{n_j \times n_c}}$$

where σ_j^2 , σ_c^2 , and σ_{jc}^2 are the variance components corresponding to the judge main effect, the case main effect and the judge-and-case interaction effect respectively; and n_c and n_j are the number of cases in the assessment and the number of judges in the Angoff procedure, respectively.

The *reliability* of the borderline regression procedure was also investigated with the RMSE. The RMSE of the cut score was calculated from the prediction error of the cut score with standard regression procedures according to the formula:

$$RMSE_{BL} = s_{reg} \sqrt{\frac{1}{n} + \frac{(R_0 - \text{mean}(R))^2}{n \times (\text{mean}(R^2) - (\text{mean}(R))^2)}}$$

where s_{reg} is the standard error of estimate of the regression of the score on the global rating; n , the sample size; R , the global rating; and R_0 , the cut point on the global rating scale.¹⁸ For comparison, the borderline-group cut-off score and the associated RMSE were also calculated. By definition, the borderline-group consists of all GPs with rating 'adequate'. The borderline-group cut-off score was obtained by calculating the average GP-score for the group of GPs with rating 'adequate', and the RMSE by calculating $SD/\sqrt{n}$, with SD the standard deviation of the GP-scores in the group of adequates, and n the number of GPs in the group.

The *perceived credibility* of the stakeholders, GPs who are assessed in their own practice, was investigated by a single-item questionnaire asking the GPs if they regarded the outcome of their individual assessments as credible. The answers were given on a five-point Likert scale, and ranged from 'completely disagree' to 'completely agree'.

RESULTS

The judges reported no problems during the Angoff and the borderline procedures. The Angoff procedure required extensive explanation before and during the procedure, but the judges felt capable of making the judgements. There was no time-order effect in the judgements.

The judges using the borderline regression method felt comfortable making even extreme judgements, though some judges would have preferred to have seven rather than five levels of competence in this procedure. The assessed GPs were classified in the Borderline regression procedure as inadequate ($n=13$), marginal ($n=18$), adequate ($n=40$), good ($n=14$) and outstanding ($n=3$).

The Angoff standard for an adequate GP was 3.1 on the MAAS-Global scale ranging from 0 to 6. The standards on the items of the MAAS-Global are given in table 1.

Table 1. The standards on the items of the MAAS-Global as obtained by the Angoff procedure.

Item	Standard	Item	Standard
Introduction	4.0	Exploration	3.0
Follow-up consultation	-	Emotions	2.3
Request for help	3.1	Information giving	3.2
Physical examination	3.6	Summarizations	2.6
Diagnosis	3.7	Structuring	3.6
Management	3.4	Empathy	3.5
Evaluation of consultation	2.1		

The borderline regression standard for an adequate GP was 2.5. The RMSE of the Angoff procedure was 0.12 on the MAAS-Global scale; the RMSE of the borderline regression procedure was 0.05. For comparison: the values for the borderline-group procedure were 2.5 with a RMSE of 0.41. An approximately double (i.e. 1.96) RMSE yields a 95% confidence interval for the test's cut score. The standards, confidence intervals and corresponding pass or fail rates are given in table 2, as are the percentages of GPs scoring within the confidence interval. For interpretation, these outcomes should be related to the scores of GPs on the video assessment; the mean scores of GPs on the MAAS-global were 2.35 ($SD\pm 0.65$).

Table 2. The standards and the confidence intervals on the MAAS-Global scale, the corresponding failure rates and the GPs scoring within the confidence interval of the Angoff, the borderline and the borderline regression method.

	Standard (CI)	Failure rate (%)	GPs scoring within CI (%)
Angoff	3.1 (± 0.2)	83	17
Borderline	2.5 (± 0.8)	62	77
Borderline regression	2.5 (± 0.1)	62	9

In the Angoff procedure, 68% of the variance in the estimated error could be attributed to the judges, 2% to the cases and 30% to the judges and the cases.

The response of the questionnaire was 38 out of 45 in the group that received feedback for the Angoff standard, and 29 out of 43 of the GPs who received feedback for the borderline regression standard. The GPs' perceived credibilities are given in table 3.

Table 3. The GPs' perceived credibility after feedback on their assessment results ($n=67$).

	Completely disagree	Disagree	Neutral	Agree	Completely agree
Angoff	2	5	12	18	1
Borderline	1	6	10	11	1

Reasons to completely disagree were not related to the standard itself, but to the perceived validity of the MAAS-Global system ($n = 2$) or to the standard-setting process itself ("GPs need support in CPD instead of standards").

DISCUSSION

Both the Angoff and the borderline regression procedure produce a credible standard according to our criteria for the assessment of doctor-patient communication in day-to-day general practice. The judges felt capable of making judgements in both procedures. The reliability of both procedures was acceptable. Although, the standards obtained by the borderline-group and borderline regression method are rather similar, the RMSE of the latter is smaller. This reflects the higher precision of the latter procedure, a finding that was expected because the borderline regression procedure uses the data of all GP's, whereas the borderline-group procedure uses only the data of the GP's in the group of adequates. The perceived credibility of the stakeholders was good. The differences in the procedures and in the outcomes, however, are striking. First, the limitations of the current study are discussed and then the stronger and weaker sides of both procedures for CPD purposes are considered.

The current study has several limitations. In the Angoff procedure, the domain is represented by eight written cases. Whether the standard can be used for samples of practising GPs' real-life cases is disputable. In test-centred methods, judgements are usually

based on the actually used items or stations, but these are not available in performance assessment. However, the cases contributed little to the variance, and therefore the use of other cases may have a limited effect on the outcome.

The failure rate is high with the Angoff standard, despite the use of a 'reality check'. One explanation is the chosen level of 'adequate performance'. However, since the goal of CPD is to sustain or improve the outcomes for the care provided, we think that no borderline performers should be excluded in the selection of GPs for CPD activities. The level of the failure rate may also be due to limitations of the procedure. To enhance the credibility of the procedure, judges were able to correct their estimations after discussion with other judges or feedback on pass or fail rates. This had a limited effect, and no time-order effect was visible. The judges may have been very experienced observers, or the complexity of the estimations may have been too great to correct them in the desired direction. Another explanation may be that in the written cases, all details about the case are given, while relevant psychological, social and medical details may remain hidden in real consultations. Aggregating estimations of items on one standard on the basis of this information may result in greater demands. Lastly, it is known that test-centred methods produce higher standards than performance-based approaches.¹⁹

In the borderline procedure, the performance level across cases were observed and judged by only one GP observer. Ideally, more observers and judges for each GP should be used, but this was not possible due to practical limitations. The reliability of the procedure was very good, but the contribution of inter-rater variability could not be estimated.

Another limitation is that the credibility of the two standards as perceived by the stakeholders was measured in different circumstances. Surprisingly, despite the large percentage of GPs with a score less than the Angoff standard requires, the majority of the GPs regarded the scores as credible, comparable to the acceptance of the lower standard obtained by the borderline regression method. This is probably due to the different feedback and setting in which the standards were presented to the GPs. Supposedly, not only the pass or fail rate, but also the setting and the use of the standard are crucial to acceptance.

As both procedures produce credible standards, preference for a specific procedure must be based on differences between both approaches, other than our used criteria. As stated, we prefer a standard that is available before assessment and gives specific feedback. Feedback is more effective when given shortly after the actual assessment. For educational activities, specific feedback regarding performance at the item level is desirable. The Angoff procedure can be carried out before assessment takes place. Further, the estimations of all judges and cases per item can be averaged to provide standards for specific feedback. Therefore, the Angoff procedure appears to be the more adequate procedure for educational purposes. The standard is high, but its main purpose is ensuring that possible under-performers are not missed when selecting GPs for CPD.

The borderline regression procedure is expensive and time-consuming if the standard is obtained in a special procedure, but no costs are involved if it is performed in routine

assessments. Therefore, with respect to feasibility, the borderline regression procedure is the procedure of first choice. If a standard has to be set for other purposes, such as re-certification, then great credibility and reliability are desired because of the far-reaching consequences for the individual GP. Time between assessment and the pass or fail decision is of less importance. The borderline regression method appears to be the desirable procedure for this purpose. It also produces standards on several levels. Not only can 'bad apples' be identified, but so can poor performers for revalidation activities and very good performers for educational tasks. For CPD purposes, however, feedback offers little specific information for improvement of specific items.

This study demonstrates that both approaches to setting standards are applicable for obtaining a standard for doctor-patient communication in daily practice. Which procedure should be chosen depends on the use of the standard. Further research to investigate the use of the procedures in other aspects of general practice is needed. Also, the possibility of selecting other judges, such as patients, for setting a standard is a challenge for further research.

LITERATURE

- Tracey JM, Arroll B, Richmond DE, Barham PM. The validity of general practitioners' self assessment of knowledge: cross sectional study. *BMJ* 1997;315:1426-8.
- Handfield-Jones RS, Mann KV, Challis ME, Hobma SO, Klass DJ, McManus IC, et al. Linking assessment to learning: a new route to quality assurance in medical practice. *Med Educ* 2002;36:949-58.
- Rethans JJ, Sturmans F, Drop R, van der Vleuten C, Hobus P. Does competence of general practitioners predict their performance? Comparison between examination setting and actual practice. *BMJ* 1991;303:1377-80.
- Ram P, Van der Vleuten C, Rethans JJ, Grol R, Aretz K. Assessment of practicing family physicians: comparison of observation in a multiple-station examination using standardized patients with observation of consultations in daily practice. *Acad Med* 1999;74:62-69.
- Zieky MJ, Livingston SA. *Manual for setting standards on the basic skills assessment tests*. Princeton, New Jersey: Educational Testing Service, 1977.
- Livingston S, Zieky M. *Passing scores: a manual for setting standards of performance on educational and occupational tests*. Princeton, 1982.
- Cusimano MD. Standard Setting in Medical Education. *Acad Med* 1996;71:S112-S120.
- Cohen AS, Kane MT, Crooks TJ. A generalized examinee-centered method for setting standards on achievement tests. *Applied Measurement in Education* 1999;12:343-66.
- Kramer AWM, Muijtjens AMM, Jansen JJM, Düsman H, Tan LH, Vleuten CPM, van der. Comparison of a rational and an empirical standard setting procedure for an OSCE. *Med Educ* 2003;37:132-9.
- Muijtjens AMM, Kramer AWM, Kaufman DM, van der Vleuten C. Using resampling to estimate the precision of an empirical standard setting method. *Applied Measurement in Education* 2003;16:245-56.
- Wilkinson TJ, Newble DI, Frampton CM. Standard setting in an objective structured clinical examination: use of global ratings of borderline performance to determine the passing score. *Med Educ* 2001;35:1043-49.
- Searle J. Defining competency - the role of standard setting. *Med Educ* 2000;34:363-66.
- Norcini JJ, Shea JA. The credibility and comparability of standards. *Applied measurement in education* 1997;10:39-59.
- Southgate L, Hays RB, Norcini J, Mulholland H, Ayers B, Woolliscroft J, et al. Setting performance standards for medical practice: a theoretical framework. *Med Educ* 2001;35:474-81.

15. Ram P, Grol R, Rethans J-J, Schouten B, Van der Vleuten C, Kester A. Assessment of general practitioners by video observation of communicative and medical performance in daily practice: issues of validity, reliability and feasibility. *Med Educ* 1999;33:447-454.
16. Van Thiel J, Ram P, Dalen Jv. *Maas-Global Manual*. Maastricht: Maastricht University, 2003.
17. Metsemakers JFM, Höppener P, Knottnerus JA, al. e. The Registration Network Family Practices: a computerized health information system in the Netherlands. *Br J Gen Pract* 1992;42:102-6.
18. Draper NR, Smith H. *Applied regression analysis*. New York: John Wiley & Sons, 1981.
19. Luijk SJv, Vleuten CPMvd. A comparison of standard setting methods applied to a performance-based test. In: Harden RM, Hart IR, Mulholland H, editors: *Approaches to the assessment of clinical competence - Part I*. Norwich: Page brothers, 1992:326-30.

Chapter 5

EFFECTIVENESS IN IMPROVING DOCTOR-PATIENT COMMUNICATION

ABSTRACT

Background

Doctor-patient communication is an essential component of general practice, and improvement of general practitioners' (GPs) communication patterns is an important target of training programmes. Available studies have so far failed to provide conclusive evidence on the effectiveness of educational interventions to improve doctor-patient communication. This chapter describes the effectiveness of the assessment-based approach to improve GPs' communication with patients.

Methods

A randomised trial was performed with a volunteer sample of 100 GPs. The intervention identified individual GPs' deficiencies in communication skills by observing authentic consultations in their own surgery. This performance assessment was followed by selection of low-scoring GPs who were offered structured activities in small-group meetings, aimed at remedying the identified shortcomings. The control group was offered existing educational materials on doctor-patient communication. The outcome measures were obtained by videotaping consultations during GPs' daily surgeries before and after the intervention. A set of 8 consultations was selected, using validated criteria and rated using the MAAS-Global.

Results

The scores in the intervention group demonstrated a significant improvement compared with those of the control group (95% CI of the difference 0.04-0.75). The effect size was moderate to large (d -value=0.66). The level of participation in the small-group meetings significantly contributed to the effectiveness. Largest improvement was found on patient centred communication skills.

Conclusion

The approach of structured individual improvement activities based on performance assessment is more effective in improving communication skills than current educational activities.

INTRODUCTION

Continuing medical education (CME) has changed considerably in the last decade as a result of theories and evidence on the effectiveness of educational activities.¹⁻⁴ In the current view, education of practicing physicians is learner-centred, focuses on day-to-day clinical practice and ideally integrates several approaches.⁵⁻⁶ This approach was adopted rapidly in many countries, even though there is conflicting evidence on its effectiveness.⁶⁻⁸

To study the effectiveness of our approach, we performed a randomised trial in general practice in The Netherlands. A subdomain of three important but very different topics was selected as an example for our study: doctor-patient communication, diabetes mellitus and ENT-disorders. We were particularly interested in the effectiveness of this approach to enhance doctors' communication skills. Doctor-patient communication is a core component of health care with a strong impact on patient satisfaction and adherence to advice and treatment.⁹⁻¹¹ Especially a patient centred approach favours these outcomes positively.¹² Although both doctors and patients put a high value on good communication, performance in day-to-day practice was found wanting.¹³⁻¹⁴ Well-designed studies of the effectiveness of educational interventions to improve communication skills are scarce and demonstrated little benefit.¹⁵⁻¹⁹ We hypothesized that a learner-centred approach, comprising individual performance assessment followed by small-group meetings tailored to doctors' individual needs might be more successful in improving doctor patient communication than traditional CME approaches.²⁰⁻²¹ A RCT was conducted to investigate the effects on general practitioners (GPs) of such a program. We also examined whether the level of participation in the intervention had an effect on the outcome, i.e. whether the more active participants achieved better results.

METHODS

Subjects

All GPs in three regions in the south of The Netherlands (n=1066) received a letter informing them about the study and asking whether they would be interested to participate. Subsequently, they were invited to enrol. We aimed at a maximum of 100 and a minimum of 80 participants. Power calculations using historical data showed these were sufficient to detect a 10% difference with 80% power and a 5% risk of a type 1 error.²² With a questionnaire participants' personal and professional characteristics (e.g. experience with modern educational methods) were obtained.

Randomisation

Randomisation was performed after sampling professional characteristics and before baseline measurement. GPs participated individually, but randomisation was done at practice level to avoid closely collaborating colleagues being randomised to different arms of the study. Practices were stratified by region and experience in working with

modern educational methods; these were identified in an expert panel prior to the study as possible effect modifiers. Subsequently balanced randomisation of practices within strata to either the intervention or the control condition was done.

Educational interventions

Intervention group

The intervention consisted of assessment, selection of global topics for improvement, feedback and revalidation activities. Assessment was the first step to identify aspects of GPs' performances for which improvement was necessary.²³⁻²⁴ Three topics were assessed. One of these topics, i.e. doctor-patient communication, is the focus of this chapter. For each participant two topics for further improvement were selected. This was done by relating assessment results to standards of desired care.²⁵ GPs who focused on other topics than communication, did not drop out of the study. The performance of GPs was assessed by video-observation in daily practice.²² Participants received individual feedback on their assessment results and the selected topics from a trained GP. Feedback was given on the participant's mean communication score in relation to the criterion-referenced standard and mean scores of 10-15 colleagues. Also, detailed feedback was on the scores on items and the standards of these items, thus enabling GPs to identify aspects requiring improvement. Within two weeks after feedback, a series of three to four small-group meetings started. The meetings lasted two hours and the groups consisted of four to six participating GPs and a tutor. The tutors were practicing regional GPs, without specific expertise in doctor-patient communication, which were prepared by an 8-hour training in supporting their colleagues' learning process. GPs also received a manual designed for the intervention, with information about quality improvement, examples, checklists, diagrams to fill out, et cetera. It contained no information about doctor-patient communication. The content of the meetings was the participants' own responsibility. GPs started by defining personal goals for improvement, then analysed barriers to change, developed practice development plans including time schedules, and finally evaluated their results.¹⁷⁻²⁶⁻²⁹ Meetings were held within a fixed time frame of seven months; participants were free to choose the time and place of the meetings. GPs received two hours of credits for performance assessment and feedback and two hours for each attended meeting.

Control group

The control group was offered existing written CME material about doctor-patient communication developed by the Dutch College of General Practitioners. It consists of paper cases, questionnaires for self-assessment and information about appropriate performance. GPs received 15 hours of credits when they returned the self-assessment questionnaires.

Variables and instruments

Doctor-patient communication

Communication skills were evaluated using videotaped consultations in GPs' own surgeries; the psychometric properties of this procedure have been described earlier.²² Observation cameras were installed in the consultation and examination room. All consultations were videotaped during 5 working days, provided the patient consented. GP observers, who received a six-hour training, selected eight consultations for each observation period using validated selection criteria.²² Observers rated communication skills using a validated instrument (MAAS-Global).^{30 31} This instrument includes twelve items on doctor-patient communication in initial consultations, rated on a scale ranging from 0 to 6; items are broadly defined but anchored in detailed rating criteria. A short description of the items is given in the *appendix*.

GPs were videotaped before and after the intervention period. The intervention group received their scores on the first observation as feedback. The control group received feedback after the study was completed. Because of a possible decay effect, the second observation was videotaped two to six months after the intervention. For the second observation the observers were blinded regarding the group to which participants belonged, i.e. intervention or control group. For the first observation blinding was not feasible, because limited time was scheduled between assessment and feedback in the intervention group.

Educational activities

The educational activities of participants were monitored using checklists and questionnaires. GPs providing feedback filled out a checklist to ensure that all relevant aspects were discussed. Tutors registered participants' attendance at group meetings. Potential bias by CME activities on doctor-patient communication besides those included in the study was identified by questionnaire.

Statistical analysis

Analyses were performed for item scores and total score. For each participant the pre and post MAAS-Global scores were calculated as the mean of all items over all observations. The item scores were calculated as the means of items over all consultations. The effect of the intervention was analysed by linear regression, using the post score as dependent variable and the pre score and treatment (intervention- or control condition) as independent variables. Possible outliers affecting the analyses were identified by means of scatter plots and excluded from analyses. To identify the contribution of the educational intervention in the post score, we calculated the 95% confidence interval of the treatment effect. Analyses were done as "intention-to-treat" analyses if data were

available: all participants, whether they participated in educational activities on doctor-patient communication or not, were included in the analysis.

The effect size (d-value) was defined on the basis of the post vs. prescores (post-pre-score) for the participants in the intervention and control group. For the two groups the average and standard deviation of the post-pre-score were calculated. The effect size was obtained by subtracting the average scores for the two groups, and dividing it by the pooled standard deviation.³²

To check whether participation in the small-group meetings influenced the effectiveness, a second regression analysis was performed. This was performed for the participants in intervention group that were allocated to the small-group meetings. In this analysis, the post score was used as dependent variable, and the pre score and the percentage of meetings attended were used as independent variables.

RESULTS

The enrolment questionnaire was returned by 670 GPs (63%); 174 (26%) indicated they were interested to participate in the study and 100 GPs, representing 87 practices, enrolled. Personal and professional characteristics were obtained; 49 participants were randomised to the intervention arm and 51 to the control arm. Flow of participants in the study and reasons for drop out are given in the flow chart.

Of the 44 GPs in the *intervention arm*, 22 were assigned to a programme of four small-group meetings about communication skills and 15 to three meetings. The remaining 7 GPs did not focus on doctor-patient communication but on other topics. They did not drop out of the study. All participants received feedback according to schedule, except one (reason: no time). Of the 37 GPs, 22 attended all meetings and four GPs did not attend any meeting. The mean number of meetings attended was 2.75. Reasons for non-attendance were: working in solo practice and being unable to attend meetings during office hours (3), and "not believing that communication training can be effective" (1). Of the GPs in the intervention group, 14 reported CME activities (mean 4.6 ± 2.7 hours) on communication not related to the study.

In the *control arm*, 16 out of 43 GPs ordered the CME materials and 16 GPs in the control group reported 6.6 ± 5.4 hours of CME on physician-patient communication not related to the study; in total 24 GPs in the control arm of the study were involved in educational activities on physician-patient communication.

The second observation was completed successfully for 38 GPs in the intervention group and 38 GPs in the control group, and used for analysis. Reasons for not participating in the second observation are given in the flow chart. GPs who did and those who did not complete the study showed no difference in scores on the first series of observed consultations (mean score 2.35 ± 0.66 vs. 2.34 ± 0.69 ; n.s.).

Over 10.000 consultations were recorded, out of which two samples of eight consultations per participant were selected. The selection criteria met in the samples were com-

parable in both observations and for both groups. The mean total scores and scores on the items of the MAAS-Global are presented in the table. No outliers in scores were identified. Regression analysis showed a significant effect of both the treatment and the prescores. The explained variance (R^2) was 0.10. The corresponding effect size (d -value) was 0.66, indicating a moderate to large effect.³² Participants reported eight items as personal improvement goals, and improvement was seen on five of these (see table). The effect of participation was analysed by a regression analysis for the participants in the intervention arm allocated to the small-group meetings, adding the percentage of meetings attended (attendance) as independent variable. In this analysis ($R^2=0.39$) the effect of the prescores ($B=0.73$; 95%CI 0.26-1.20) and the effect of the attendance ($B=1.30$; 95%CI 0.57-2.04) were found to be significant.

Table. Item and total scores of GPs in intervention ($n=38$) and control group ($n=38$), and the 95% CI of the differences in improvement between both groups.

	Intervention		Control		Difference
	Pre $\pm$ SD	Post $\pm$ SD	Pre $\pm$ SD	Post $\pm$ SD	95% CI
MAAS-Global					
item scores ¹					
Introduction	2.81 $\pm$ 0.80	3.11 $\pm$ 0.92	3.01 $\pm$ 0.81	2.64 $\pm$ 0.98	0.09 - 0.95
Request for help	0.85 $\pm$ 0.82	1.53 $\pm$ 1.14	1.21 $\pm$ 1.07	0.69 $\pm$ 0.75	0.47 - 1.35
Physical					
examination	3.24 $\pm$ 0.98	3.75 $\pm$ 1.13	3.45 $\pm$ 0.97	3.75 $\pm$ 1.02	-0.44 - 0.57
Diagnosis	3.08 $\pm$ 0.70	3.42 $\pm$ 0.99	3.18 $\pm$ 0.90	3.14 $\pm$ 0.92	-0.10 - 0.74
Management	2.69 $\pm$ 0.89	3.18 $\pm$ 1.24	2.94 $\pm$ 0.86	3.03 $\pm$ 0.98	-0.27 - 0.74
Evaluation of consultation	1.17 $\pm$ 1.00	1.34 $\pm$ 1.06	1.28 $\pm$ 1.12	0.69 $\pm$ 1.04	0.15 - 1.11
Exploration	1.78 $\pm$ 0.72	2.14 $\pm$ 1.30	1.93 $\pm$ 0.90	1.53 $\pm$ 1.17	0.16 - 1.23
Emotions	0.83 $\pm$ 0.65	0.78 $\pm$ 0.90	0.99 $\pm$ 0.98	0.39 $\pm$ 0.58	0.03 - 0.73
Information giving	2.94 $\pm$ 0.82	3.20 $\pm$ 0.91	3.28 $\pm$ 0.90	3.16 $\pm$ 0.82	-0.29 - 0.51
Summarizations	1.01 $\pm$ 0.81	0.99 $\pm$ 1.11	1.73 $\pm$ 1.23	0.80 $\pm$ 0.99	-0.30 - 0.72
Structuring	3.22 $\pm$ 1.02	3.32 $\pm$ 1.11	3.35 $\pm$ 0.91	3.41 $\pm$ 1.01	-0.52 - 0.42
Empathy	3.15 $\pm$ 1.12	3.87 $\pm$ 1.44	3.56 $\pm$ 1.21	3.57 $\pm$ 1.36	-0.31 - 1.00
MAAS-Global					
total scores	2.22 $\pm$ 0.54	2.53 $\pm$ 0.82	2.48 $\pm$ 0.73	2.21 $\pm$ 0.72	0.04 - 0.75

¹ The bold items were reported by participants as personal improvement goals.

Figure. Flow chart of participants and reasons for dropout. **Reasons for dropout:**

- A. No time (n=4), no reason (n=1)
- B. No time (n=8), no reason (n=1)
- C. No time (n=1), technical problems (n=1)
- D. No time (n=1), illness (n=1), leaving primary care (n=1), not recording an adequate number of consultations (n=1).
- E. Technical problems (n=1)
- F. Not recording an adequate number of consultations (n=1), technical problems (n=1), no time (n=1)

DISCUSSION

The educational approach involving assessment of communication in daily practice and personalised learning activities guided by structured small-group work proved to be an effective way of improving doctor-patient communication. The size of the effect was moderate to large. A dose-response effect was seen as the attendance to the intervention significantly contributed to the improvement of the scores. The participants showed most improvement on patient-centred items.

Participants in the intervention group undertook more study-related activities to improve doctor-patient communication than their colleagues in the control group. This may be interpreted as bias, but in our view it should be regarded as an effect of the intervention, i.e. the assessments, to drive learning activities. Educational activities not related to the study may have influenced the results, but – as GPs in the control arm attended more of those activities than those in the intervention arm – this only emphasizes the effectiveness of the intervention. Selection of randomised groups by attrition may cause bias; although we cannot rule out this effect, no differences in available prescores were found between those who did and did not drop out of the study.

The effects of the intervention may be temporary and fade over time. Therefore, we planned the second measurement two to six months after the intervention. Whether improvements are permanently embedded in GPs' communication behaviour is not clear.

This study showed assessment of GPs performance in the real process of care delivery to be an effective first step of an educational programme. GPs considered the assessment outcomes as an acceptable basis for educational activities, and selected aspects of their performance with low scores for improvement. GPs were also found to be capable of drawing up effective personal development plans with support only for the process of change, given the improvement that was seen on the reported improvement goals. The participating GPs had to get used to the approach investigated in this study, as the majority reported to have no experience with the methods used in the intervention. We expect that repeated use of the programme may yet enhance its effectiveness.

The findings of this study cannot easily be generalised to other topics or to the general population of GPs. The participants in our study were not representative as they were motivated to experience this new approach. Effectiveness may be different when GPs are not really interested in this educational method. Also, the assessment method of video-observation in daily practice is ideal for this purpose, as it is personalised and performance-based and yields videotapes that can be used in the small-group meetings. Further research will have to shed light on the effectiveness of this approach for other topics, assessment formats and participating doctors.

LITERATURE

1. Knowles MS. *Self-directed learning. A guide for learners and teachers*. Cambridge: The Adult Education Company, 1975.
2. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing Physician Performance. *JAMA* 1995;274:700-705.
3. Davis D, O'Brien MA, Freemantle N, Wolf FM, Mazmanian P, Taylor Vaisey A. Impact of formal continuing medical education: do conferences, workshops, rounds, and other traditional continuing education activities change physician behavior or health care outcomes? *JAMA* 1999;282:867-74.
4. Oxman AD, Thomson MA, Davis DA, Haynes RB. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *CMAJ* 1995;153:1423-1431.
5. Thomson O'Brien MA, Freemantle N, Oxman AD, Wolf F, Davis DA, Herrin J. Continuing education meetings and workshops: effects on professional practice and health care outcomes. *Cochrane Library* 2001:1-42.
6. Cantillon P, Jones R. Does continuing medical education in general practice make a difference? *BMJ* 1999;318:1276-9.
7. Peck C, McCall M, McLaren B, Rotem T. Continuing medical education and continuing professional development: international comparisons. *BMJ* 2000;320:432-5.
8. Grol R. Improving the Quality of Medical Care: building bridges among professional pride, payer profit and patient satisfaction. *JAMA* 2001;286:2578-85.
9. Salmon P, Sharma N, Valori R, Bellenger N. Patients' intentions in primary care: relationship to physical and psychological symptoms, and their perception by general practitioners. *Soc Sci Med* 1994;38:585-92.
10. Williams S, Weinman J, Dale J. Doctor-patient communication and patient satisfaction: a review. *Fam Pract* 1998;15:480-92.
11. Levinson W, Roter DL, Mullooly JP, Dull VT, Frankel RM. Physician-patient communication. The relationship with malpractice claims among primary care physicians and surgeons. *JAMA* 1997;277:553-7.
12. Little P, Everitt H, Williamson I, Warner G, Moore M, Gould C, et al. Observational study of effect of patient centredness and positive approach on outcomes of general practice consultations. *BMJ* 2001;323:908-11.
13. Grol R, Wensing M, Mainz J, Ferreira P, Hearnshaw H, Hjortdahl P, et al. Patients' priorities with respect to general practice care: an international comparison. European Task Force on Patient Evaluations of General Practice (EUROPEP). *Fam Pract* 1999;16:4-11.
14. van den Brink Muinen A, Verhaak PF, Bensing JM, Bahrs O, Deveugele M, Gask L, et al. Doctor-patient communication in different European health care systems: relevance and performance from the patients' perspective. *Patient education and counseling* 2000;39:115-27.
15. Hutchinson L. Evaluating and researching the effectiveness of educational interventions. *BMJ* 1999;318:1267-9.
16. Torgerson CJ. Educational research and randomised trials. *Med Educ* 2002;36:1002-1003.
17. Hulsman RL, Ros WJ, Winnubst JA, Bensing JM. Teaching clinically experienced physicians communication skills. A review of evaluation studies. *Med Educ* 1999;33:655-68.
18. Cegala DJ, Lenzmeier Broz S. Physician communication skills training: a review of theoretical backgrounds, objectives and skills. *Med Educ* 2002;36:1004-16.
19. Ihler E. Patient-Physician communication. *JAMA* 2003;289:92.
20. Epstein RM, Hundert EM. Defining and assessing professional competence. *JAMA* 2002;287:226-235.
21. Hobma SO, Ram PM, Merode GGv, Vleuten CPM, van der, Grol RPTM. Feasibility, costs and appreciation of a tailored approach for continuing professional development of general practitioners. *Quality in Primary Care* 2004;12:271-8.
22. Ram P, Grol R, Rethans J-J, Schouten B, Van der Vleuten C, Kester A. Assessment of general practitioners by video observation of communicative and medical performance in daily practice: issues of validity, reliability and feasibility. *Med Educ* 1999;33:447-454.
23. Tracey JM, Arroll B, Richmond DE, Barham PM. The validity of general practitioners' self assessment of knowledge: cross sectional study. *BMJ* 1997;315:1426-8.

24. Handfield-Jones RS, Mann KV, Challis ME, Hobma SO, Klass DJ, McManus IC, et al. Linking assessment to learning: a new route to quality assurance in medical practice. *Med Educ* 2002;36:949-58.
25. Hobma S, Ram P, Muijtjens AMM, Grol R, van der Vleuten C. Setting a standard for performance assessment of doctor-patient communication in general practice. *Med Educ* 2004;38:1244-52.
26. Bennett NL, Davis DA, Easterling WE, Friedmann P, Green JS, Koeppen BM, et al. Continuing medical education: a new vision of the professional development of physicians. *Acad Med* 2000;75:1167-72.
27. Berwick DM. Continuous improvement as an ideal in health care. *N Engl J Med* 1989;320:53-6.
28. Berwick DM. A primer on leading the improvement of systems. *BMJ* 1996;312:619-622.
29. Batalden PB, Stoltz PK. A framework for the continual improvement of health care: building and applying professional and improvement knowledge to test changes in daily work. *Jt Comm J Qual Improv* 1993;19:424-47.
30. Van Thiel J, Ram P, Dalen Jv. *Maas-Global Manual*. Maastricht: Maastricht University, 2003.
31. http://www.hag.unimaas.nl/Maas-global_2000/.
32. Colliver JA. Effectiveness of problem-based learning curricula: research and theory. *Acad Med* 2000;75:259-66.

Chapter 6

EFFECTIVENESS IN IMPROVING CLINICAL PERFORMANCE ON ENT-DISORDERS

ABSTRACT

Background

ENT-disorders are frequently presented problems in general practice and several guidelines that describe appropriate care are available. In daily practice, the adherence to these guidelines shows a large variation. An important problem is the overprescription of antibiotics, while the recommendation of guidelines is that most patients do not need these. Previous studies have not provided conclusive evidence for effective strategies to change clinical performance. This study investigates the effectiveness of the assessment-based approach to improve the care for ENT-disorders.

Methods

A randomised trial was performed in general practices in The Netherlands. In total 42 practices participated in the control group, and 45 practices in the intervention arm. The main outcome measures were indicators on structure and process, derived from general practice guidelines; these were sampled by (1) questionnaires on practice organisation and equipment (2) GPs' prospective recordings of performance and (3) prescription data.

Results

A small but non-significant improvement on primary effect measures was found comparing the intervention group with the control group. A trend for a decrease in the number of antibiotic prescriptions was found in the intervention arm; due to a significant decrease in the number of prescriptions of antibiotics that are advised against in the guidelines. Subgroup analyses showed no effect of the invested time in the intervention.

Conclusions

Our intervention did not demonstrate effectiveness in improving relevant aspects of care for ENT-disorders. The disappointing result can partly be explained by shortcomings in our study, such as insufficient power or insensitivity to the effect measures, or by a lack of face-validity of the knowledge test used. Probably, the topic (ENT-disorders) to which GPs were not highly motivated is an important determinant for effectiveness. We therefore question the applicability of assessment-based education as a generic approach for practising physicians.

INTRODUCTION

Over the last decades, several strategies to improve quality of care have been developed.¹ According to current insights in continuing professional development, learning of professionals should actively involve the learner in the learning process, point at the objective learning needs and be tailored to the specific needs and situations.²⁻⁸ Despite high expectations, the evidence on the effectiveness of existing approaches is not straightforward.^{19 10}

We developed a programme for physicians' education that fits in with the current view on learning of professionals.^{5 11} It is an individually tailored programme that starts with a needs-assessment to select aspects of performance in need of improvement.^{12 13} On the basis of assessment results, a selection out of the assessed subdomain is made for improvement. Participants consequently meet up in small groups to make individualised development and improvement plans.

To study the effectiveness of our approach, we performed a randomised trial in general practice in The Netherlands. A subdomain of three important but very different topics was selected as an example for our study: doctor-patient communication, diabetes mellitus and ENT-disorders. The intervention appeared to be effective in improving doctor-patient communication.¹⁴ This chapter describes the effects on ENT-disorders.

ENT-disorders were chosen as these represent common problems in general practice on which seven clinical guidelines are now available in The Netherlands. Guidelines have been introduced to support physicians in clinical decision-making, but despite obvious advances, implementation of guidelines into clinical practice has proved to be complex.¹⁵⁻¹⁷ The adherence to ENT-guidelines also showed flaws, like a considerable over prescription of antibiotics (Spies, unpublished data). In previous studies, particularly the over prescription of antibiotics has shown to be resistant to change.^{18 19}

The study question is whether our approach that starts with selecting participants and topics for change and subsequently offers a structured quality improvement program in small-group meetings is more effective in changing clinical performance than current educational approaches.

METHODS

A cluster-randomised trial was performed with GPs in the intervention group receiving the developed approach and GPs in the control arm receiving usual CME. The approach was studied using the available guidelines on ENT-disorders in The Netherlands: acute otitis media, otitis media with effusion, otitis externa, sore throat, allergic and hyper reactive rhinitis and impaired hearing.

Subjects

We aimed at a minimum of 80 and a maximum of 100 participating GPs; power calculations using historical data showed this was enough to detect a 10% difference in adher-

ence to guidelines between intervention and control group with 80% power and a 5% risk of a type 1 error. After subscription, personal and professional characteristics of participants were collected by questionnaire.

Randomisation

Randomisation was done after sampling the professional characteristics. The intervention focused on individual GPs. However, randomisation and analyses were done at practice-level. This was done to avoid that closely collaborating colleagues were distributed over the two study-arms. Before randomisation, practices were stratified on the basis of three characteristics: (1) region, (2) experience with the educational techniques used in the intervention, and (3) the use of a computer aided prescription system, which was introduced nationwide shortly before the study and which supports GPs in writing prescriptions according to the guidelines. The strata were randomised by computerised allocation to the intervention or control condition.

Educational interventions

Intervention group

The intervention consisted of a CPD-programme with assessment, feedback and a personalised educational programme for improvement supported by small-group meetings.

Assessment. Assessment was the first step to select areas in need of improvement. Three topics were assessed. ENT disorders were assessed by a written knowledge test, which provides a feasible alternative to performance assessment.^{20 21} Criterion-referenced minimum standards, obtained by Angoff-procedures, were available for all assessments.^{22 23} A preset procedure selected for each participant two topics for improvement activities, linking assessment results to criterion-referenced standards or, in case of indecisive results, by following GP's preferences.

Feedback. Participants got individual feedback on their assessment results and selected topics. This consisted of written materials, presenting the participant's mean scores, the criterion-referenced standards and scores of 10-15 colleagues. Furthermore, individual scores and standards on specific ENT disorders were given, to allow GPs to identify specific improvement goals. A trained GP-colleague visited participants in their own practice to present these materials and advise on preparations for the first small-group meeting.

Small-group meetings. Within two weeks, participants started with a series (3-4) of 2-hour meetings in a small group of 4-6 GPs. A trained GP-tutor and a manual supported GPs in the process of improving through a detailed, step by step programme aimed at daily practice. The manual provided examples, checklists, diagrams to be completed, et cetera, but no information on ENT-disorders. Participants determined the content of the

meetings themselves, thus tailoring them to their specific needs.²⁴ The meetings were held within a fixed time frame of 7 months; participants could also tailor the exact time and place to their demands.

Educational materials

To support the medical contents of the meetings, guidelines and related educational materials of the Dutch College of General Practitioners were available, or participants could consult experts or order handbooks.

Credit points

As an incentive, participants received 2 credit points for the assessment and feedback and for every meeting attended.

Control group

The GPs were offered existing educational materials about ENT-disorders, developed by the Dutch College of General Practitioners. These materials contain paper cases, questionnaires for self-assessment, information on appropriate performance and materials for small-group education.

Variables and instruments

The effect of the intervention was determined with a triangulation of methods, focusing on aspects of structure and process.²⁵ Data were collected before and after the educational interventions.

1. *Structure: practice organisation.* Aspects of good practice organisation are explicitly described in guidelines, or else guidelines implicate the availability of equipment and adjustment of practice organisation; this was investigated by means of a questionnaire. The questions concerned different dimensions of organising the care for ENT-disorders.
2. *Process: performance in consultations.* This was measured by prospective recording of performance. GPs received six recording forms containing key questions on the clinical case and the treatment for the three most frequent ENT-disorders: otitis media, otitis media with effusion and otitis externa ENT-disorders. They were asked to complete a recording form after consultations concerning these conditions. This method reveals more relevant information to judge guideline adherence than chart audit or direct observation.²⁶
3. *Process: prescriptions.* Combining data of the regions' two main health-insurers, data on more than 70% of GPs' prescriptions were available. Indicators on prescriptions regarding ENT-disorders were determined over a period of one year before and one year after the study. Indicators were derived from guidelines. Also, the DU90%, a simple quantitative method for assessing the quality of drug prescribing, was determined. It expresses the number of different products prescribed, accounting for 90% of the prescriptions within a specific ATC-category.²⁷

Educational activities

The actual participation in intervention activities was investigated using checklists. The feedback-giving GP filled in a checklist to test if all relevant aspects had been discussed. Tutors registered the presence of participants in the group meetings. Participants in both study-arms were questioned about their planned improvements in their daily care for ENT-disorders. They were also asked if they had conducted CME-activities not related to the study ENT-disorders during the study period, to detect possible bias.

Statistical analysis

Variables for structure were determined per practice by dividing the aspects of practice organisation in conformity with the guidelines by the maximum possible score. Scores were calculated for all subdimensions as well as overall scores. Variables for process were determined per practice by dividing all decisions in conformity with the guidelines by the total number of measured decisions. Scores were calculated for each disease separately as well as an overall ENT-score. Adherence to recommendations on prescriptions were calculated as percentages of adherence, in which the nominator is the number of recommended prescriptions and the denominator is the number of all prescriptions in the specific therapeutic group (ranging for 0-100%, with a higher score indicating better performance). DU90% is presented in absolute numbers (with a lower score indicating a better performance). Besides these relative indicators, absolute numbers of prescriptions on antibiotics were determined, considering that over-prescription of antibiotics is the main problem in general practice. Total scores for prescriptions were administered by adding all available relative indicators on a scale ranging from 0-1 with the inverse DU90%.

The effect of the intervention was analysed using linear regression, using the post score as dependent variable and the following independent variables: the pre score, intervention or control condition, time since GP registration, practice type (single-handed, two GPs, group practice, health centre), availability of a practice nurse, availability of a computer aided prescription system, region, and experience with educational techniques used in the intervention. These variables were indicated as possible effect modifiers in an expert meeting prior to the study. Variables that did not attribute to the explained variance were removed from the model. Practice organisation, data from self-recording and prescribing behaviour were analysed separately. Both total scores and specific dimensions were analysed. Analyses were done as 'intention-to-treat' analyses if data were available.

To check if participation in the intervention influenced the effectiveness, a separate analysis of the participants in the intervention group that were allocated to the small-group meetings on ENT-disorders was done, taking the same dependent and independent variables and adding as an independent variable the percentage of planned meetings attended.

RESULTS

We recruited 100 GPs, representing 87 practices. Professional characteristics were collected. Practices were randomised to the intervention (n=45) or control (n=42) condition.

During baseline measurements, GPs left the study because of a lack of time (5) or without apparent reason (2). During the intervention period 4 GPs stopped for the following reasons: no time (2), illness (1) and leaving the profession (1). This limited both groups to 38 practices with 44 GPs in the control group and 45 GPs in the intervention arm. Relevant characteristics are given in table 1.

Table 1. Characteristics of analysed practices and participating GPs.

		Intervention (n=38)	Control (n=38)
Practice type:	single-handed	12	14
	partnership	24	20
	health centre	2	4
Availability of a practice nurse		4	2
Availability of a computer aided prescription system		11	7
Region:	A	14	13
	B	10	10
	C	14	15
GPs experienced in educational methods used in the intervention		17	18
Time since registration (years $\pm$ SD)		13.7 $\pm$ 7.2	14.1 $\pm$ 7.3

In the *intervention arm*, 39 GPs (representing 35 practices) were allocated to the small-group meetings about ENT-disorders (17 for 4 meetings and 22 for 3 meetings) on the basis of their assessment results. All participants but one (reason: no time) got feedback according to plan. Of this group, 12 GPs attended all meetings, and 5 did not attend any meetings. The mean number of meetings attended was 2.1. Reasons for not participating were practical grounds or a lack of interest in the subject. Reported goals for improvement of active participants were: reduction in number of antibiotic prescriptions (11), changing preferences for ENT-medication (8), improving diagnostic skills and equipment (7) and reduction of number of referrals (2). GPs in the intervention arm reported a mean of 0.6 (SD $\pm$ 1.8 hours) of accredited extra CME-activities on ENT-disorders not related to the study.

In the *control arm* 23 practices ordered the CME-materials on ENT-disorders. No educational activities were reported to obtain accreditation points. They reported a mean of 0.4 $\pm$ 1.0 hours of accredited CME-activities on ENT-disorders not related to the study, which was not statistically different from GPs in the intervention arm.

The practices in the intervention group showed a slightly larger, but non-significant improvement in practice organisation compared to practices in the control condition (table 2). Data on self-recording were available for 536 consultations on ENT-disorders, representing 3151 clinical decisions (table 3). More increase in adherence was seen in the intervention group, but differences were not statistically significant.

Data on prescriptions (table 4) showed a non-significant effect on the total number of prescribed antibiotics ($p=0.06$), due to a decrease in the number of less preferable antibiotics prescriptions in the intervention arm ($p<0.05$). Other variables showed no significant differences between the intervention and control group.

Table 2. Aspects of practice organisation of participating practices for ENT-disorders, as reported by questionnaire (n=number of practices; mean score $\pm$ SD).

Aspect (max score)	Intervention		Control	
	Pre (n=38)	Post (n=36)	Pre (n=36)	Post (n=36)
Total score (29)	11.19 $\pm$ 3.87	12.80 $\pm$ 4.05	10.88 $\pm$ 4.38	11.90 $\pm$ 4.00
Available equipment:				
Medical technical (8)	3.57 $\pm$ 1.40	3.78 $\pm$ 1.42	3.36 $\pm$ 1.56	3.59 $\pm$ 1.52
Patient education (6)	3.73 $\pm$ 2.09	4.35 $\pm$ 1.90	3.33 $\pm$ 2.16	3.93 $\pm$ 2.11
Use of equipment (11)	3.70 $\pm$ 1.95	4.42 $\pm$ 1.92	3.93 $\pm$ 2.00	4.18 $\pm$ 2.05
Quality improvement (4)	0.09 $\pm$ 0.25	0.12 $\pm$ 0.29	0.13 $\pm$ 0.31	0.10 $\pm$ 0.27

Table 3. Adherence to recommendations of guidelines, obtained by prospective recording of participants during daily surgeries. The percentages (mean $\pm$ SD) of decisions taken according to the guideline are given (n=number of practices).

	Intervention		Control	
	Pre (n=30)	Post (n=22)	Pre (n=30)	Post (n=27)
Overall	64 $\pm$ 11	70 $\pm$ 12	65 $\pm$ 13	65 $\pm$ 17
Otitis externa	57 $\pm$ 15	68 $\pm$ 17	57 $\pm$ 19	59 $\pm$ 22
Acute otitis media	64 $\pm$ 17	65 $\pm$ 19	68 $\pm$ 19	63 $\pm$ 23
Otitis media with effusion	77 $\pm$ 18	81 $\pm$ 13	74 $\pm$ 16	79 $\pm$ 14

Table 4. Indicators (mean $\pm$ SD) on prescriptions of participating practices regarding ENT-disorders, determined during one year before and during one year after the educational interventions (n=number of practices).

	Intervention		Control	
	Pre (n=35)	Post (n=35)	Pre (n=33)	Post (n=33)
Antibiotics:				
DU90%	4.26 $\pm$ 1.29	4.40 $\pm$ 1.19	3.94 $\pm$ 1.59	4.23 $\pm$ 1.15
All prescribed antibiotics (prescriptions/year)	443.6 $\pm$ 204.3	466.6 $\pm$ 210.9	397.6 $\pm$ 225.5	421.6 $\pm$ 203.3
Advised antibiotics (%)	59.4 $\pm$ 9.8	58.5 $\pm$ 9.3	61.3 $\pm$ 9.8	55.9 $\pm$ 9.1
Ear drops:				
Advised ear drops (%)	24.4 $\pm$ 20.6	31.0 $\pm$ 21.3	27.7 $\pm$ 20.7	32.3 $\pm$ 24.4
Anti allergies :				
DU90%	9.00 $\pm$ 2.01	7.95 $\pm$ 2.71	8.77 $\pm$ 1.98	8.17 $\pm$ 2.40

The proposed model explained 0% variance in self-recording, 17% in prescription data and 60% in practice organisation (adjusted R^2). Beside the effects on antibiotic prescriptions, the intervention did not contribute significantly to the explained variance. Of the other independent variables, only the pre-measurements contributed significantly to the variance. Subgroup-analysis of the intervention group showed no significant effect of the number of meetings attended.

DISCUSSION

The intervention showed a small, non-significant effect on the quality of care for ENT-disorders, as compared to the control condition. With respect to the prescription of antibiotics an improvement was found on one aspect, namely a reduction of less preferable antibiotics. Several factors may cause the lack of convincing positive results.

Aspects in the study itself may only partly explain the absence of positive results. A weakness in our design is inherent to a prospective study of a personalised assessment-based approach. Participants have selected topics for improvement themselves in conformity with the developed approach.²⁵⁸ Effects of the intervention may have remained invisible as effect measures had to be designated prior to the study and therefore may be different from the actual improvement. For example, one participant decided to improve her skills in otoscopy, which was not measured in our study. Also, the assumed 10% improvement in power calculation may have been too optimistic. Despite these weaknesses, limited relevant improvement in relevant performance was found.

Specific aspects of our intervention could account for the disappointing result. Similar studies found no effects of assessment-based approaches either.²⁸²⁹ One possible explanation was that assessment should be more specific.²⁸ We think however that the assessment used in our study focused on well-defined aspects of general practice. Perhaps assessment of actual performance, instead of knowledge, should be a starting

point.³⁰ The face-validity of the knowledge test may have been insufficient and not motivating for improvement activities. Also, the selection of GPs and topics may have been less accurate than necessary given the fact that the relation between knowledge and performance is not straightforward.^{31 32} Previous research however appointed knowledge tests as a valid and feasible alternative to performance assessment.^{20 21}

The small-group meetings following the assessments did not affect the performance of GPs in the intervention group either, although this is a method with a reasonable chance to be effective.^{6 7 33} As described in the previous chapter, these meetings contributed to the effectiveness of the intervention in changing doctor-patient communication.¹⁴ Therefore, we think that the lack of effectiveness cannot be contributed to materials or methods used, but to other aspects in the study.

What are these other factors that are responsible for the lack of effectiveness? In evaluating the intervention, several participants and tutors stated that ENT-disorders are neither inspiring nor very important in daily care. Also, the attendance to the meetings, the adherence to the program and the appreciation of the participants was different, compared to other topics.³⁴ Therefore, we hypothesise that the topic of interest was a major factor that has influenced effectiveness.^{35 36} Although ENT-disorders have a high prevalence in general practice, the results of the knowledge test were poor, our data on performance showed considerable variation between practices as well as ample room for improvement, and despite the fact that over prescription of antibiotics is a serious problem, our participants appeared to be not very motivated to improve.

Literature about assessment of competence describes a large variation in scores between different cases for the same candidates, known as case-specificity.^{35 37} If this can be partly explained by motivation for specific topics, and low scores indicate a lack of motivation, educational approaches like ours, that rely on motivation, cannot be expected to be effective. The approach towards education of practising physicians, based on objective weaker points in delivered care, may therefore depend on this important factor.

Our main conclusion is that we could not demonstrate the superiority of an assessment-based approach, focusing on real needs of practising physicians. Recently, the need of needs assessments has been emphasised by several authors.³⁴ Our study indicates that a needs assessment may be necessary to determine learning needs, but the next step towards actual improvement of care may be problematic. Further research is needed to determine the influence of specific factors like the attitudes of GPs.³⁵ In our view, this should focus on the question if the effectiveness of assessment-based approaches can be improved by assessing topics that are 'interesting for physicians to improve', or by enhancing 'the interest for physicians to improve' e.g. by not relying on internal professional motivation but by for example financial incentives or criterion-referenced standards for recertification.

LITERATURE

1. Grol R. Improving the Quality of Medical Care: building bridges among professional pride, payer profit and patient satisfaction. *JAMA* 2001;286:2578-85.
2. Grant J. *The Good CPD Guide*. Sutton: Reed Business Information, 1999.
3. Grant J. Learning needs assessment: assessing the need. *BMJ* 2002;324:156-9.
4. Norman GR, Shannon SI, Marrin ML. The need for needs assessment in continuing medical education. *BMJ* 2004;328:999-1001.
5. Bennett NL, Davis DA, Easterling WE, Friedmann P, Green JS, Koeppen BM, et al. Continuing medical education: a new vision of the professional development of physicians. *Acad Med* 2000;75:1167-72.
6. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing Physician Performance. *JAMA* 1995;274:700-705.
7. Davis D, O'Brien MA, Freemantle N, Wolf FM, Mazmanian P, Taylor Vaisey A. Impact of formal continuing medical education: do conferences, workshops, rounds, and other traditional continuing education activities change physician behavior or health care outcomes? *JAMA* 1999;282:867-74.
8. Knowles MS. *Self-directed learning. A guide for learners and teachers*. Cambridge: The Adult Education Company, 1975.
9. Beyer M, Gerlach FM, Flies U, Grol R, Król Z, Munck A, et al. The development of quality circles/peer review groups as a method of quality improvement in Europe. *Fam Pract* 2003;20:443-51.
10. Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health technology assessment Winchester, England* 2004;8:iii-iv, 1-72.
11. Cantillon P, Jones R. Does continuing medical education in general practice make a difference? *BMJ* 1999;318:1276-9.
12. Tracey JM, Arroll B, Richmond DE, Barham PM. The validity of general practitioners' self assessment of knowledge: cross sectional study. *BMJ* 1997;315:1426-8.
13. Handfield-Jones RS, Mann KV, Challis ME, Hobma SO, Klass DJ, McManus IC, et al. Linking assessment to learning: a new route to quality assurance in medical practice. *Med Educ* 2002;36:949-58.
14. Hobma SO, Ram PM, Muijtjens AMM, Vleuten CPM, van der Grol RPTM. Effective improvement of doctor-patient communication: a randomised trial. *Accepted for publication by the British Journal of General Practice*.
15. Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. *Lancet* 1993;342:1317-22.
16. Freemantle N, Harvey EL, Wolf F, Grimshaw JM, Grilli R, Bero LA. Printed educational materials: effects on professional practice and health care outcomes. *The Cochrane Library* 2001:1-24.
17. Feder G, Eccles M, Grol R, Griffiths C, Grimshaw J. Clinical guidelines: using clinical guidelines. *BMJ* 1999;318:728-30.
18. Flottorp S, Havelsrud K, Oxman AD. Process evaluation of a cluster randomized trial of tailored interventions to implement guidelines in primary care--why is it so hard to change practice? *Fam Pract* 2003;20:333-9.
19. Flottorp S, Oxman AD, Havelsrud K, Treweek S, Herrin J. Cluster randomised controlled trial of tailored interventions to improve the management of urinary tract infections in women and sore throat. *BMJ* 2002;325:367.
20. Ram P, Vleuten Cvd, Rethans J, Schouten B, Hobma S, Grol R. Assessment in general practice: the predictive value of written-knowledge tests and a multiple-station examination for actual medical performance in daily practice. *Med Educ* 1999;33:197-203.
21. Aeschlimann A, Westkaemper R, Doherty M, Woolf AD. Multiple choice question quiz: a valid test for needs assessment in CME in rheumatology and for self assessment. *Annals of the rheumatic diseases* 2001;60:740-3.
22. Cusimano MD. Standard Setting in Medical Education. *Acad Med* 1996;71:S112-S120.
23. Hobma S, Ram P, Muijtjens AMM, Grol R, van der Vleuten C. Setting a standard for performance assessment of doctor-patient communication in general practice. *Med Educ* 2004;38:1244-52.

24. Campbell SM, Sheaff R, Sibbald B, Marshall MN, Pickard S, Gask L, et al. Implementing clinical governance in English primary care groups/trust: reconciling quality improvement and quality assurance. *Qual Saf Health Care* 2002;11:9-14.
25. Donabedian A. *Explorations in quality assessment and monitoring 1. The definition of quality and approaches to its assessment.*, 1980.
26. Spies TH, Mokkink HG, De Vries Robbe PF, Grol RP. Which data source in clinical performance assessment? A pilot study comparing self-recording with patient records and observation. *Int J Qual Health Care* 2004;16:65-72.
27. Bergman U, Popa C, Tomson Y, Wettermark B, Einarson TR, Aberg H, et al. Drug utilization 90%—a simple method for assessing the quality of drug prescribing. *Eur J Clin Pharmacol* 1998;54:113-8.
28. Borgiel AE, Williams JJ, Davis DA, Dunn EV, Hobbs N, Hutchison B, et al. Evaluating the effectiveness of 2 educational interventions in family practice. *CMAJ* 1999;161:965-70.
29. Wright J, Warren E, Reeves J, Bibby J, Harrison S, Dowsell G, et al. Effectiveness of multifaceted implementation of guidelines in primary care. *Journal of Health Research & Policy* 2003;8:142-8.
30. Dean B. Learning from prescribing errors. *Qual Saf Health Care* 2002;11:258-60.
31. Rethans JJ, Sturmans F, Drop R, van der Vleuten C, Hobus P. Does competence of general practitioners predict their performance? Comparison between examination setting and actual practice. *BMJ* 1991;303:1377-80.
32. Ram P, Van der Vleuten C, Rethans JJ, Grol R, Aretz K. Assessment of practicing family physicians: comparison of observation in a multiple-station examination using standardized patients with observation of consultations in daily practice. *Acad Med* 1999;74:62-69.
33. Oxman AD, Thomson MA, Davis DA, Haynes RB. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *CMAJ* 1995;153:1423-1431.
34. Hobma SO, Ram PM, Merode GGv, Vleuten CPM, van der, Grol RPTM. Feasibility, costs and appreciation of a tailored approach for continuing professional development of general practitioners. *Quality in Primary Care* 2004;12:271-8.
35. Watkins C, Harvey I, Carthy P, Moore L, Robinson E, Brawn R. Attitudes and behaviour of general practitioners and their prescribing costs: a national cross sectional survey. *Qual Saf Health Care* 2003;12:29-34.
36. Van Berkel HJM, Schmidt HG. Motivation to commit oneself as a determinant of achievement in problem-based learning. *Higher Education* 2000;40:231-42.
37. Wass V, Van der Vleuten C, Shatzer J, Jones R. Assessment of clinical competence. *Lancet* 2001;357:945-9.

Chapter 7

FEASIBILITY AND APPRECIATION

ABSTRACT

Background

There is a growing tendency to develop more complex interventions for continuing professional development (CPD) of physicians in order to enhance effectiveness. Besides their effectiveness, it is interventions' feasibility and the appreciation of stakeholders that are increasingly regarded as key features for their implementation in daily educational routines. The objective of this study is to investigate the feasibility and appreciation of such a complex approach to CPD in which GPs are assessed and subsequently work in small groups on the improvement of individual demonstrated deficiencies.

Methods

A cohort study was performed in general practices in The Netherlands with 43 volunteer participants. The main outcome measures were the ability of GPs and supporting staff to perform the intervention; the costs per hour; and the participants' appreciation of (aspects of) the educational intervention.

Results

GPs are able to perform a CPD-intervention that starts with a needs assessment and that subsequently supports the individual self-directed learning process. GPs need on average 22.3 hours for the assessments, small-group meetings and work in their practices. Costs are € 117.56 per hour. The mean appreciation is 6.8 on a 10-point scale. Appreciation of and participation in the intervention are dependent on the topic studied.

Conclusions

The approach towards continuing professional development is feasible and acceptable. It requires a context in which sufficient resources are available with respect to budget, educational materials and skilled support staff. Furthermore, GPs must be really interested in the topic studied and probably also in the specific approach.

INTRODUCTION

According to current insights, continuing professional development (CPD) of practising physicians should be multifaceted, self-directed, learner-centred and should focus on the actual needs of physicians or health care organisations.¹⁻⁸ Translating these demands into an educational intervention requires some form of assessment of current competence or performance, preceding the intervention, to begin with. The assessment results must then be used to define improvement goals, which may differ for each person or health care organisation. Learning plans should be pointed at these specific goals and they should be individualised, depending on differences in personal or local situations and needs. This educational ideal places great demands on both learners and organisers of CPD, since both individual learning paths for physicians and combining several educational methods enhance the complexity of educational interventions.

In order to implement these interventions successfully, they should be effective, feasible, and physicians must be willing to participate. We studied these aspects in an educational intervention for general practitioners, based on the principles described above. The intervention consisted of assessments, feedback and a personalised programme aiming at improvement of actually delivered care. This chapter focuses on the feasibility and appreciation of the intervention. Feasibility is defined as the extent to which participants and involved support staff could perform the intervention as planned, and as time and budget required. Appreciation is defined as the value of the intervention to the participating GPs. We were interested to know what factors related to these primary effect measures. Therefore, we also studied various aspects of the intervention that we expected to be of influence, such as the instruments and procedures used, the contents of the programme and the acceptance of the underlying guidelines.

METHODS

Subjects

A letter informed all GPs in the south of The Netherlands ($n=1066$) about the study; they were asked if they were willing to participate in our study. We aimed at a minimum of 40 participants in the intervention. The questionnaire was returned by 670 GPs (63%), out of who 174 (26%) showed interest in participating, and a maximum of 100 actually subscribed. This chapter focuses on the GPs that were randomised to the intervention arm of the study and who participated actively in the intervention ($n=43$).

Educational intervention

The intervention (see flow chart) consisted of assessments, feedback and a personalised programme of self-directed learning, aiming at improvement of daily care. We used *assessments* to select aspects of care in need of improvement. Three topics were used as examples in the project to study our approach. Doctor-patient communication, which was assessed by video-observation during daily surgeries,⁹ and competence in manage-

ment of Diabetes Mellitus and of ENT-disorders, assessed by written knowledge tests. By relating assessment results to predefined standards, we selected two topics for each GP.^{10 11} Following this procedure, GPs were allocated to doctor-patient communication (n=37), ENT-disorders (n=39) and Diabetes Mellitus (n=10). Participants got individual *feedback* on their test results. They received a feedback-report with their mean scores, the predefined standards and scores of 10-15 colleagues, including detailed feedback on sub items, to allow identification of specific personal improvement goals. A GP-colleague, who had attended a three-hour training for this specific purpose, visited participants in their own practices, elucidated written feedback and gave instructions to prepare the first small-group meeting. Subsequently, participants could attend a series of 7 *small-group meetings* (4-6 GPs) of 2 hours each. Trained GP-tutors and a manual supported them through a step-by-step programme aiming at improvement of daily practice. Participants first defined an individual goal for improvement. On the basis of subsequently determined barriers, they made personal development plans; these were readjusted if necessary and were evaluated in the second part of the programme. Support focussed on the process of learning and improving; the manual provided examples, checklists, diagrams to fill in, etcetera, but contained no specific information about the clinical contents. Tutors were practising GPs, prepared for their task by an 8-hour training.

Support staff

A total of 50 collaborators were involved in the intervention (see table 1).

Educational materials

Educational materials on the topics, e.g. textbooks and written materials from the Dutch College of General Practitioners based on current guidelines, were supplied if necessary and support was offered in case of specific needs, e.g. a meeting with an expert.

Figure. Flowchart of the educational intervention.

Time schedule

Assessments were planned over a maximum period of three months and feedback within the following two weeks.¹² Seven months were allowed for the small-group meetings that started in the week after the feedback; participants could tailor these to their demands. Assessments, feedback and the first small-group meetings were planned within a limited timeframe to enhance effectiveness.^{12 13} This implicated that assessment results of all participants had to be available at the end of the planned assessment period to allow distribution of the participants over small groups.

Credits

As an incentive, 4 credit points were offered for the assessments and feedback and 2 credit points for each meeting attended. GPs need 40 credit points per year for recertification.

Table 1. Aspects of the intervention, educational activities of the participating GPs, required support staff and their tasks, and instruments and materials needed.

Aspect of intervention	Instruments & materials	Support staff	Task of support staff
Assessment			
Doctor-patient communication	Video-equipment	Secretary	Planning and distribution of materials
	Predefined standard	Technician	Installing video-observation systems and instructing GPs
		Participant	Making recordings/logbook
		GP-observer	Selecting and observing consultations
ENT-disorders	Written knowledge test	Participant	Doing the test
	Predefined standard		
Diabetes Mellitus	Written knowledge test	Participant	Doing the test
	Predefined standard		
Allocation of GPs to topics		Researchers	Allocating GPs to topics in need of improvement on basis of assessment results
Feedback	Software to convert scores into written feedback	Secretary	Production of written feedback
		GP giving feedback	Elucidating written feedback & preparing 1 st small-group meeting
Small-group working	Manual for participants	GP-tutors	Presiding small-group meetings
	Manual for tutors		
Overall organisation		Secretary/ researchers	Planning and monitoring

Dependent variables

Feasibility was operationalised firstly as the ability of participating GPs and support staff to perform the intervention as planned, secondly as the costs related to the intervention and thirdly as the appreciation of the intervention by the participants.

Performing the intervention

To answer the question if the participants and support staff were able to perform the intervention as planned, we investigated the participation in the intervention, i.e. the extent to which participants completed the programme and were able to do this within the scheduled timeframe, and the number of drop-outs. As criteria were not available in literature, we defined the following three criteria for feasibility: (a) a majority (>50%) of the participating GPs had to be able to complete all stages of the intervention including self-reported improvement on one or two topics; (b) a limited drop out of participants due to organisational problems within the planned time schedule (max 5%); (c) a maximum assessment-and-feedback time of 4 months and a period of 8 months between feedback and the last small-group meeting.

Costs

These were determined as the required investment in Euros per hour for every participant. The costs are running costs and are determined in an ideal size unit for implementation, whereas the variation in costs in other sizes was calculated.

Costs were calculated on the basis of the following. The nominator included the costs of the participant, the time invested by support staff, and the costs of instruments and materials needed in the intervention. All time dedicated to the intervention (assessment, feedback, small-group work and improvement activities in own practice) was included in the denominator. The ideal sample size was assumed to be a unit of 6 GPs. Variations in costs in other sample sizes were calculated by adding or removing 2 GPs, assuming shifts in costs due to needs in terms of video-equipment, GP-tutors etcetera.

GP-fees (participants, feedback giving GPs, GP-tutors, observers of video consultations and researchers) were assumed to correspond to an hourly rate of € 77. -.¹⁴ The costs of the participants are opportunity costs, as GPs cannot deliver patient care during educational activities; costs for the support staff are true costs. Secretarial costs were calculated on the basis of the current salary scales (€13. -/hr). Fees for installers of video-equipment were calculated on the basis of the true costs (€18. -/hr).

The costs of all instruments and materials needed were calculated using the annuity method, assuming a depreciation period of 5 years for all instruments and materials except the knowledge tests and the accompanying standards, for which a period of 3 years was assumed. The costs were determined assuming a running period of 1 year with normal use of the instruments and materials.

Appreciation

Participants' appreciation of the intervention as a whole was investigated by questionnaire and rated on a scale ranging from 1 to 10, the traditional Dutch scoring scale in education. We regarded a score of 7 or higher as satisfactory, and a score of 5 or lower as insufficient.

The appreciation of various aspects of the intervention was given on a 5-point Likert scale, ranging from 1 (=dislike) to 5 (=high appreciation). These aspects were related to

the intervention, i.e. the assessment methods used, the predefined standards of desired performance, the procedure of allocating GPs to topics on the basis of assessment results, the written and personal feedback, and the support in the small-group meetings. Other aspects were related to the clinical contents that were used to study our intervention, i.e. the perceived relevance of the topic for improvement and the agreement with the existing guidelines on the topics used.

Instruments

Questionnaire

Participating GPs completed a questionnaire shortly after the intervention. This questionnaire contained a checklist to investigate the adherence to the programme in the small-group meetings. It also contained open questions concerning the required time-investment for participation in the intervention as well as its general appreciation, and closed questions on the appreciation of various aspects of the intervention. Besides these questions, participants were invited to comment on the intervention and specific aspects.

Adherence to the time-schedule

We carefully monitored and registered the performance of each aspect of the intervention up to the first small-group meeting, and compared these to the initial time-schedule.

Time-registration

All support staff prospectively registered the time required for their tasks to determine the costs of the intervention.

Statistical analysis

Analyses were done using SPSS. Differences between appreciations of different aspects of the intervention (the assessment methods, the predefined standards, the allocation to the selected topics, the perceived relevance of the topic for improvement and the agreement with the existing guidelines on the topics used) were analysed by T-tests.

The relation between the appreciation of various parts of the intervention (the assessment methods used, the predefined standards of desired performance, the procedure of allocating GPs to topics on the basis of assessment results, the written and personal feedback, the support in the small-group meetings, the perceived relevance of the topic for improvement and the agreement with the existing guidelines on the topics used) and participation was investigated by correlating the appreciations with the presence or absence of self-reported improvement.

RESULTS

The questionnaire was returned by 42 of the 43 GPs. Data on all participants' time-schedules and complete time-registrations of the support staff were available.

Feasibility

Performing the intervention

Of the 42 participants, 14 reported to have successfully completed the programme on two topics; 18 reported completion on one topic, indicating that the majority (74%) were able to perform the approach. There was no dropout of the study due to problems in the organisation. The planned period of three months for the assessments and feedback was feasible with regard to the written knowledge tests, but in the video-assessment 20 participants exceeded the 4 months; 11 needed one extra week, the remaining 9 GPs exceeded it with 2-6 weeks, mainly due to absence of GPs or increased workload due to absence of colleagues over the holidays. The planned period of 7 months for the small-group meetings was sufficient for all groups. Table 2 shows the participation in the different steps of the intervention.

Table 2. Self-reported steps taken by GP-participants in the CPD program.

Number of GPs	Doctor-patient communication	ENT-disorders	Diabetes Mellitus	Overall
Allocated to the topic	37	39	10	86
Defining a goal for improvement	32	30	8	70
Making a personal development plan	26	27	8	61
Achieving goal for improvement	23	15	8	46

Costs

Besides the time dedicated to assessment, feedback and small-group meetings (mean 13.1 ± 4.4 hours), the participants reported to have invested a mean of 9.2 ± 9.9 hours for implementing the improvement plans into daily practice. Costs per hour are given in table 3. The costs are € 117.56 per hour with a proportion of 65% of opportunity costs. Diminishing or enlarging the assumed ideal unit of 6 GPs with 2 GPs increases the costs with € 8.07 (6,8%).

Table 3. Costs of (aspects of) the intervention per hour per participant in a unit of 6 GPs and costs in a unit of 6±2 GPs.

Aspect of intervention	Support staff/instruments/materials	Costs/hour	
		6 GPs	6±2 GPs
	Participant ¹	€ 77,-	-
Assessment	<i>Doctor-patient communication</i>		
	Video-equipment ²	€ 3.29	€ 4.94
	Secretary ²	€ 0.10	-
	Installer ²	€ 1.21	-
	GP-observer ²	€ 12.66	-
	Predefined standard ²	€ 0.16	-
	<i>ENT-disorders</i>		
	Written knowledge test ²	€ 1.51	-
	Predefined standard ²	€ 1.04	-
	<i>Diabetes Mellitus type 2</i>		
	Written knowledge test ²	€ 0.68	-
	Predefined standard ²	€ 0.39	-
	<i>Allocation of GP to topics</i> ¹	€ 0.58	-
Feedback	Software ¹	€ 0.56	-
	Production of written feedback ¹	€ 0.44	-
	Giving oral feedback ¹	€ 3.39	€ 3.67
Small-group working	GP-tutors ¹	€ 12.27	€ 18.41
	Manual for participants ¹	€ 0.45	-
	Manual for tutors ¹	€ 0.10	-
Overall organisation	Planning and monitoring ¹	€ 1.73	-
Total costs/hour		€ 117.56	€ 125.63
Total costs/GP		€ 2609.10	€ 2801.55

1 = fixed costs 2 = variable costs

Appreciation

The mean appreciation of the intervention as a whole on a 10-point scale was 6.8; 30 GPs (70%) gave an appreciation of 7 or higher and 4 GPs (9%) gave an appreciation of a 5 or lower. Four participants did not give a score for the approach as a whole, as we asked them to do, but made a distinction between topics; these GPs appreciated doctor-patient communication more than ENT-disorders (8.0 vs. 5.0).

The appreciation of various parts of the intervention was investigated and possible differences and effects on the participation were analysed.

Assessment

The video assessment of actual performance was appreciated more than the knowledge tests with regard to the identification of stronger and weaker sides (3.93 vs. 3.34; $p < 0.05$), being an incentive to improve (3.86 vs. 3.08; $p < 0.05$) and a means to select topics for improvement (3.98 vs. 3.30; $p < 0.05$). There were no differences in appreciation between the two knowledge tests on these aspects. Despite the substandard scores on doctor-patient communication and ENT-disorders, criterion-referenced standards were accepted by the participants. There were no differences between the assessments (3.62 vs. 3.71 vs. 3.29). Participants reported a neutral attitude towards the selection of topics for improvement on the basis of the assessment results. Participants in the groups on doctor patient communication were however more satisfied about their allocation than those on ENT-disorders (4.09 vs. 3.57; $p < 0.05$).

Feedback

The participants appreciated written feedback positively. The oral explanation by a GP-colleague was considered to be non-threatening but to contribute little to the information given on paper. Some participants would have preferred feedback from the GP who had done the video-observations.

Small-group meetings

Both the tutor and the manual were evaluated positively; the role of the tutor was described as indispensable. As many as 27 participants planned to further use the approach of the small-group meetings, focusing on the process of improvement, in the future, while 8 definitely did not want to experience this educational approach again. Their reasons were that it is too laborious, too time-consuming while, as a participant stated, "CME should be relaxing". Two participants explicitly reported that the CPD-approach was too time-consuming; the total of 7 small-group meetings was too much, according to 22 participants. The remaining GPs regarded the time needed for the approach as acceptable. Several participants complained that the amount of credit points was not proportionate to the hours invested.

Besides the instruments and procedures, the clinical contents of the meetings and the underlying guidelines may be of influence. Participants reported ENT-disorders to be

less appropriate for our approach, compared to doctor patient communication ($p < 0.05$). The acceptance of existing guidelines was comparable for all topics. Various parts of the intervention were expected to influence the participation in the intervention. We found a significant correlation between the extent to which the intervention was performed and the perceived suitability of our approach for the topic of interest. This was found in both ENT-disorders ($r = -0.41; p < 0.05$) and doctor-patient communication ($r = -0.45; p < 0.01$); there were not enough data on diabetes for these calculations. The opinions of participants on other aspects studied, such as the assessments used and the attitude towards the content of the guidelines involved, showed no significant correlations.

DISCUSSION

The intervention is feasible when assessed with our criteria, costs are €117.- to €125.- per hour (adding up to a mean of €2700.- per participant) and a majority of 70% of our participants valued the approach positively. However, our findings must be interpreted carefully and we will therefore discuss the three aspects in more detail.

Our approach appeared to be feasible as 74% of our participants successfully completed the intervention. This is an important finding as it shows the ability to pass through a self-directed learning trajectory based on weaknesses revealed by assessments. In our study, a group of doctors interested in this approach was studied in only two learning cycles. Repeated use of the approach may enhance this number of completed learning cycles. Our study was, however, done in a selected group of GPs, interested in our approach. The capability of doctors, less interested in this approach, may be different, as the approach relies heavily on the individual activities of the participating doctors. Therefore, to understand more about the feasibility to implement our intervention in educational routines for practising doctors, more research is needed.

With regard to costs we deliberately did not formulate explicit criteria for acceptability, nor did we perform a cost-benefit analysis. As we intended our approach to be generically applicable for all subdomains of general practice, both the costs of the assessments and the benefits will differ strongly depending on the topic. Also, these must be weighed considering the local situation. Whether costs are acceptable depends on the budget available and on the resources already spent on education. In our study, 65% of the costs are opportunity costs. These opportunity costs are small if compared to many popular educational conferences, in which GPs are out of practice for days, or sometimes a week. In the latter, the costs are 40-50 hours, while they receive education for no more than 20-30 hours. Furthermore, an important part (69%) of the running costs are fees for GP-colleagues, who themselves may also benefit from being involved. These effects were however outside the scope of our study.

To successfully implement the intervention in daily practice, the micro-economy of the doctors in the target group must be considered. A number of participants complained about the time-investment, inside and outside their practice, and about the reward in

terms of credit points. The benefits for participating doctors, in terms of work-satisfaction, credit-points or money, must be made clear, also for hours spent in practice, as these diminish the costs per hour and are necessary for effectiveness.

The majority of our participants appreciated the approach investigated, and only a small minority assessed the intervention negatively. These are the opinions of a selected group. Before recruiting participants for the study, we found 26% in the GP-population interested in our approach. Whether this reflects a large group with limited interest or a promising group of early adopters is not clear.¹⁵

We found considerable variation in the appreciation and performance of the intervention between the topics studied, as described in the previous chapters. This conflicts with our aim to develop a generically applicable approach for all subdomains of general practice. Assessments were used to detect areas of care in need of improvement, as doctors tend to follow courses on topics they are already good at.¹⁶⁻¹⁸ Our findings suggest that the approach used, may not be the final solution for this problem.⁵

LITERATURE

1. Knowles MS. *Self-directed learning. A guide for learners and teachers*. Cambridge: The Adult Education Company, 1975.
2. Fox RD, Bennett NL. Learning and change: implications for continuing medical education. *BMJ* 1998;316:466-8.
3. Bennett NL, Davis DA, Easterling WE, Friedmann P, Green JS, Koeppen BM, et al. Continuing medical education: a new vision of the professional development of physicians. *Acad Med* 2000;75:1167-72.
4. Grant J. *The Good CPD Guide*. Sutton: Reed Business Information, 1999.
5. Handfield-Jones RS, Mann KV, Challis ME, Hobma SO, Klass DJ, McManus IC, et al. Linking assessment to learning: a new route to quality assurance in medical practice. *Med Educ* 2002;36:949-58.
6. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing Physician Performance. *JAMA* 1995;274:700-705.
7. Oxman AD, Thomson MA, Davis DA, Haynes RB. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *CMAJ* 1995;153:1423-1431.
8. Davis D, Evans M, Jadad A, Perrier L, Rath D, Ryan D, et al. The case for knowledge translation: shortening the journey from evidence to effect. *BMJ* 2003;327:33-5.
9. Ram P, Grol R, Rethans J-J, Schouten B, Van der Vleuten C, Kester A. Assessment of general practitioners by video observation of communicative and medical performance in daily practice: issues of validity, reliability and feasibility. *Med Educ* 1999;33:447-454.
10. Cusimano MD. Standard Setting in Medical Education. *Acad Med* 1996;71:S112-S120.
11. Hobma S, Ram P, Muijtjens AMM, Grol R, van der Vleuten C. Setting a standard for performance assessment of doctor-patient communication in general practice. *Med Educ* 2004;38:1244-52.
12. Brinko KT. The Practice of Giving Feedback to Improve Teaching. *Journal of Higher Education* 1993;64:574-592.
13. Rolfe I, McPherson J. Formative assessment: how am I doing? *Lancet* 1995;345.
14. Verstappen WJHM, Merode Fv, Grol R, Grimshaw J, Dubois W, Weijden Tvd. Comparing Cost Effects Of Two Quality Strategies to Improve Test Ordering in Primary Care. A Randomized Trial. *International Journal of Quality in Health Care* 2003;16:391-8.
15. Rogers EM. *Diffusion of innovations*. 3rd ed. New York: The Free Press, 1983.
16. Fox RD, Harvill LM. Self-assessments of need, relevance and motivation to learn as indicators of participation in continuing medical education. *Med Educ* 1984;18:275-81.
17. Fox RD, Mazmanian PE, Putnam RW. *Changing and learning in the lives of physicians*. New York: Praeger, 1989.

18. Tracey JM, Arroll B, Richmond DE, Barham PM. The validity of general practitioners' self assessment of knowledge: cross sectional study. *BMJ* 1997;315:1426-8.

Chapter 8

GENERAL DISCUSSION

Introduction

This chapter deals with the conclusions and implications of the study. First, the findings and conclusions with regard to the research questions are summarised. Then, the developed concept of 'directed self-learning' and the intervention are discussed. Subsequently, possible flaws in the study with regard to the conclusions are considered. Finally, we focus on the practical implications of the study for the continuing education of medical professionals and for research on continuing professional development (CPD).

Findings and conclusions

The aims of the study were to develop an intervention to support CPD and compare its effectiveness and feasibility with current routines in continuing medical education. More specifically, the research questions were: (1) what are the features of an approach that can be expected to support CPD of medical professionals effectively and efficiently; (2) is this approach more effective in changing GPs performances towards desired standards than current educational approaches; (3) is this approach feasible, i.e. what are the needs in terms of time and money and do GPs appreciate the developed approach? Below, the findings are discussed per question.

Concept & intervention

An approach of 'directed self-learning' was proposed. It was hypothesised that a doctor's educational process must be partly 'directed', and should partly rely on 'self-learning'. This was worked out into an intervention, using the results of a study on the needs and barriers as reported by the target group, practising Dutch GPs. A pilot study was performed to examine and further improve aspects of the intervention's feasibility. Based on theoretical considerations and practical experiences, the following intervention was designed for further study. The first step was to make assessments, in order to support GPs in identifying aspects of care in need of improvement ('directed learning'). To select topics for improvement, assessment results were linked to predefined criterion-referenced standards of adequate performance. GPs subsequently received feedback on their results, as well as the topics selected for improvement. After this, they met up in small groups, guided by a tutor. In these groups, GPs were supported in the process of performance-improvement ('directed learning'), but the contents and targets of their activities were their own individual and professional responsibility ('self-learning').

Effectiveness

The effectiveness (i.e. changing GPs' performances towards desired behaviour) of this approach was compared to current educational approaches in a randomised trial. Three topics out of the broad domain of general practice were selected for the study: doctor-patient communication, ear-, nose and throat (ENT) disorders and type 2 diabetes mellitus.

The effectiveness found, turned out to differ between the topics studied. These results allow no straightforward answer to the question regarding the intervention's effectiveness. Besides the educational approach, the topic of interest appears to be an important factor in the effectiveness of an intervention. This conflicts with the aim to develop a generically applicable and effective approach.

Doctor-patient communication was studied by observing GPs' authentic consultations, videotaped during daily surgeries before and after the study. The scores in the intervention group demonstrated a significant improvement compared with those of the control group, with a moderate to large effect size. The level of participation in the small-group meetings significantly contributed to the effectiveness. Doctor-patient communication was the topic that most GPs preferred for the small-group meetings, and tutors reported inspiring meetings with motivated GPs.

The effectiveness with respect to ENT-disorders was studied with a triangulation of different effect measures, derived from general practice guidelines. Learning needs were selected, using a written knowledge test. Differences in improvement between the two randomised groups were not found. Subgroup analyses showed no effect of the time invested in the intervention. The motivation to change performance appeared to be limited, as only one GP preferred ENT-disorders after the assessments, attendance of the small-group meetings was less compared to the other topics, and GP-tutors presiding the small groups on ENT-disorders reported that a lot of participants were difficult to motivate, if at all possible.

Only a small minority of GPs in the intervention was allocated to type 2 diabetes mellitus, as assessment results revealed less learning needs compared to the other topics. Given the number of participants improving their performance in diabetes, and the concomitant loss of power, the results regarding this topic were not described.

Feasibility

Feasibility (i.e. the needs in terms of time and money) and GPs' appreciation of the developed approach was studied in the pilot and, more detailed, in the main study. The costs per participant were € 2,600.- to € 2,800.- with a proportion of 65% opportunity costs due to the time dedicated to the intervention by the participants themselves. Participants invested on average 22.3 hours in the intervention. Participants appreciated the intervention with a mean score of 6.8 on a 10-point scale. Appreciation of the intervention appeared to depend on the topic studied.

These results indicate the intervention is feasible in the context of general practice in The Netherlands. But feasibility is not an isolated characteristic of an intervention; it must be linked to its effectiveness and other aspects of the context, e.g. the support and resources that are available. The conclusion on the intervention's possible role in CPD must be determined by weighing both effectiveness and feasibility together.

Conclusion

The aim was to develop a generic approach that supports GPs' CPD more effectively and efficiently than current approaches. Resuming the results, it proved to be difficult to develop such an approach. Taking into account the ambiguity of its effectiveness, the considerable costs in terms of time and money, and the moderate enthusiasm of the participants, we need to conclude that the study does not provide evidence that the intervention, based on the concept of directed self-learning, is superior to current educational support of GPs. Whether the results must be attributed to the concept of directed self-learning, to the intervention itself, or to flaws in the study, is gone into below.

Concept & intervention

Concept

'Directed self-learning' is a theoretically sound and defensible concept but, in a more generic use, it may not fit in well with the current context of general practice. The assumption that objective needs (as revealed by assessments) direct and motivate GPs towards desired performance, is too optimistic. Learning and improving in the current context is still directed by individual GPs' motivations and perceived needs.

The approach studied appears to be more effective than current approaches under specific conditions, that is if perceived and objective needs are congruent. If this is the case, the approach provides a more effective alternative for current routines. However, motivation makes or breaks the approach: if it lacks, as appeared to be the case in ENT-disorders, the approach can be expected to be ineffective.

To reconsider the strengths and weaknesses of the concept's crucial elements in more detail, three aspects will be discussed: identifying the learning needs (by assessment), supporting the process of changing performance (by small-group working) and the individual professional's autonomy (the participants' responsibility for their own learning trajectories).

Learning needs. With regard to assessment, there is a tension between necessity and effectiveness. It is necessary to identify objective learning goals, but problems arise in the effectiveness of assessment in initiating improvement activities.^{1,2}

The study confirms the need of assessment to select priorities for learning agendas. Due to the assessments, participants' activities in the intervention arm focused more on actual needs, compared to those in the control arm. Assessments successfully directed participants towards aspects of their communication patterns to be improved, and the knowledge tests appointed ENT-disorders as a learning priority, contrary to GPs' preferences. However, in the latter case, GPs were not led ('directed') towards effective improvement activities. This tension between necessity and effectiveness was found previously in a similar approach, and in our preceding studies.³⁻⁵

Supporting change. Supporting participants in the process of improving care is necessary and can be effective. Stepwise quality improvement is not a method that is effective in itself, it requires the proper context and knowledge.^{6 7 8} Effectiveness of the approach depends on the available equipment and routines. This can be compared to care provision, for example a surgeons' equipment and experience are major determinants for the success of a treatment.⁹⁻¹² This also goes for the effectiveness of learning. Participants need 'improvement-knowledge' and 'improvement-skills', or else should be supported in these aspects of professional learning. The participating GPs recognised this, and valued the tutor's role highly, but they accepted support only if they were interested to improve their performance.

Professional responsibility. The participant's professional responsibility to initiate and complete improvement activities is, like the role of assessment, both necessary and problematic. It is necessary since 'professional development' per se implies the interference of the professional. It is problematic because professionals have certain autonomy in deciding what they want to learn or improve, and this may conflict with the intention to direct participants towards desired performance. If participants do not agree with the identified learning needs, as appeared to be the problem for a number of participants with regard to ENT-disorders, the approach will be ineffective as it leans heavily on the individual's intrinsic professional motivation.

The apparent absence of interest in improving ENT-disorders reflects a professional attitude towards learning that differs from our implicit basic assumptions. Support in CPD can only be effective if it corresponds with the professional habits and values of the target group. We implicitly assumed a commitment to achieve the highest possible standards in patient care.^{13 14} The approach requires learners who are directed and motivated by something as abstract as 'objective learning needs' or 'demonstrated deficiencies'. However, this was falsified by the study results.

Participants appeared to be mainly interested in the specific content of small-group working and their perceived possible gains. This fits in a professional culture that emphasises individual professional autonomy, with self-directed learning as an associated learning style. This type of professionalism is more and more criticized in literature, and as it is a possible major effect modifier in CPD, it is discussed in more detail under 'implications for CPD and further research'.

Intervention

It is not only the concept, but also many other factors that influence the effectiveness of educational interventions: 'the devil is in the details'.^{15 16} In translating the concept into the actual intervention, several pragmatic choices and compromises on aspects of the intervention had to be made. These may very well have influenced the results. Four apparently important aspects are discussed, being the assessments used, the allocation of participants to topics, the feedback and the generic support in the small groups.

Assessments. The assessments used in the intervention did not come forward to the demands that had been defined. Despite more than 15 years of developing assessments for general practice, a problem was encountered in the availability of appropriate assessment methods. In ENT-disorders and diabetes, assessment was therefore done using knowledge tests, while we preferred assessment of performance.

However, the use of a knowledge test alone cannot be held responsible for the disappointing results in the case of ENT-disorders. It may be stated that the poor results of assessments combined with the fact that ENT-disorders are very common disorders, should have stimulated GPs to develop improvement activities.

Allocation of participants. We chose a procedure of allocating participants to topics primarily on the basis of assessment results, which may have been too straightforward. Assessment can be combined with other educational activities in many different ways.³

¹⁷⁻²¹ A possible and effective approach is to restrict assessments and educational activities to motivated doctors.^{17 18} This is however a concession to the ideal of improving objective needs, as it allows individual deficiencies to sustain. Therefore we do not regard this as an appropriate alternative for our approach.

Feedback. A better balance needs to be found between the effectiveness and the feasibility of our approach in order to provide participants with feedback on the assessment results. To enhance the effectiveness, a considerable investment was made to meet current insights.²² Feedback was provided shortly after the assessments, in addition, software was developed to present scores in a visually attractive way, and finally, a group of GPs was specially trained to visit all participants individually in their own practices.

Despite these efforts, participants were not very enthusiastic. This may be explained by many factors, as feedback is a pivoting point in our intervention. Appreciation may for example very well be influenced by the fact that GPs are not used to giving or receiving feedback in this very personal way, or do not feel confident about the assessments or the assessment results. Weighing the investments and the outcomes in terms of effectiveness and appreciation, we think improvements should primarily be pointed at approaches to feedback that are more feasible to organise and better accepted by participants.

Generic support. We offered generic support in the small-group meetings, instead of content-specific support in the topic as is usually done in continuing education. This has some clear advantages but also a few disadvantages. As it is completely open, it offers the possibility for educational activities to be tailored exactly to the participants' learning needs and context; it is also more feasible as specific support requires more in terms of learning materials or specialists' expertise. However, in a generic approach, developing expertise and collecting learning materials is a burden that is shifted on the shoulders of the participants, who often have limited experience and competencies in these educational matters.

More specific support in the topic, i.e. materials and programs that are ready to use, has advantages in feasibility from the participants' view. Also, the unique individuality and need of personal tailoring must not be overestimated, as similar learning needs and barriers towards improvement are often found.²³⁻²⁶ From an organisational view however, this requires a much larger infrastructure in terms of materials, specialist-tutors etcetera, because learning needs cannot be predicted before the assessment. More support in the content of a topic in small-group meetings may be a less feasible, but a more effective alternative to our intervention. Although relevant, the choice for generic support alone cannot be held responsible for our results, given the positive results in doctor-patient communication.

Flaws in the study

What aspects in the study may limit the validity of the conclusions concerning the concept of directed self-learning? Aspects related to study power, to effect measures and to the population studied will be discussed.

Power

We compared the effectiveness of two educational approaches that differed in a number of aspects. One essential difference between the approaches in the intervention arm and the control arm was the way GPs selected topics for improvement, i.e. assessment-based vs. self-directed selection. This had to be involved in the study design as it is an essential element of the contrast between the two approaches. The result is however that for each topic only a subgroup of participants with an unpredictable size is actively involved in educational activities. In analysing the effectiveness, the possible effects of the active subgroup are diluted over all participants entering the study. In diabetes, this resulted in a considerable loss of power, but in the other two topics studied, large groups of GPs participated in study activities, and allowed analysis of the results.

Effect measures

The effects on ENT-disorders were studied by a triangulation of methods already available; in the case of practice organisation and prescribing behaviour, the specific instruments were adapted for this study. In all three effect-measures, the sensitivity to measure change was not studied previously. However, it is to be doubted if this can explain the results. Most improvement goals were defined in the subdomain of altering prescribing behaviour. Despite extensive information on prescriptions, differences between both study arms could not be found.

External validity

The external validity of the findings is limited by the fact that we studied a selected group of GPs, who were interested to experience the directed self-learning approach. Although we did not have any problems in recruiting GPs for the study, the majority of GPs was not interested in the intervention itself. We may hypothesise that effectiveness

requires GPs who are interested in both the topic and the educational approach, and therefore the approach will be less effective in a random group of GPs.

Implications for CPD

As professionals play a key role in health care, optimising their performance is essential to sustain and improve the quality of care. Therefore, continuing education of the health care professional remains indispensable, and we need effective approaches to support health care professionals in their continuing development. The challenge is to optimise both effectiveness and feasibility of approaches to support CPD.

Developments in CPD must aim at both a conceptual level and a practical level. On a conceptual level, the focus must be on the coherence between the professional culture and educational support in CPD. On a practical level, educational instruments and needs, for example assessments or expertise in stepwise quality improvement, must be further developed. First, the development of educational instruments and needs will be discussed. Then, we will focus on the educational concept and the professional culture.

Educational needs

To perform and study new approaches to CPD, educational instruments and expertise are required. At the start of the study, we encountered problems in the availability of educational instruments such as assessments and specific expertise, for example GPs to cooperate in the intervention (as a tutor, to give feedback or to observe videotaped consultations). Regardless of new concepts and approaches to CPD, some necessary educational instruments and expertise that must be further developed can be pointed out. We will discuss these necessary requirements in supporting CPD, as well as the improvements needed.

Assessments are necessary and must be further developed. They should aim at the level of performance. Theoretically, performance assessment is essentially different from competence based performance, and also GPs prefer these assessments.^{5 27-29} Assessment results should inspire GPs to start learning and should be easy to translate into defining goals for improvement. This requires that realistic and acceptable standards of adequate performance be defined.

Assessment implies feedback on assessment results; the feasibility of giving feedback as well as the effectiveness of using feedback for improvement must be enhanced. Feasibility may for example be improved by using new technical developments, like web cams that allow a limited number of experienced collaborators to give personal, non-threatening feedback without wasting time on travelling. Effectiveness may be improved by enhancing the experience of GPs in receiving feedback, e.g. by giving it on a regular basis.³⁰ Furthermore, the effectiveness of feedback is of course also affected by the assessments used and the consequent learning format.

More knowledge and skills on the process of change are necessary, and an infrastructure of 'improvement-knowledge', e.g. well-trained tutors, must be established. Participants described the tutors' role as very positive, but we question if the tutors with their current

training could contribute enough to the meetings. In evaluating the study with the GP-tutors, several problems were reported by them. Thorough training is necessary and the 8-hour training that tutors received in our study must be regarded as a minimum. Besides training in generic educational knowledge and skills, content-specific support is needed.^{15 31} Maybe, tutors' training should also comprise this, as content-specific expertise may influence the effectiveness of problem-based learning groups in a positive way.¹⁶ Also, more directive content-specific materials or methods to support small-group working can be used.³² Whether this provides an approach in which the benefits in terms of effectiveness weigh against the costs in terms of a loss in feasibility has to be further studied.

The effectiveness of learning approaches is not based on learning methods alone, but also on the people using them. The effects of new approaches to CPD may be enhanced by enlarging the expertise in using new methods by the target group as discussed above, but also by increasing the awareness of current problems in learning and. It is necessary to provide GPs with general information about the problems in current CME, and about the opportunities of new approaches. During the recruitment of participants for our study, one of the major reasons for not participating was the fact that GPs reported to be satisfied with their current learning activities. In my personal experience, a number of false assumptions on learning must be removed before GPs show interest in other educational approaches. Therefore, the target group must understand the ideas and facts behind the approach. This may also contribute to achieve necessary changes in the professional culture towards learning, which is one aspect of the next discussion.

Professional culture

We concluded that the developed approach to CPD contains some necessary elements but, as a whole, it does not fit in with the professional culture. To discuss this, we first go back to where we started in chapter 2, the definition of CPD. We proposed the following definition: "continuing professional development is a doctor's ongoing learning process in order to meet medical and societal standards of desired performance in patient care."

CPD is primarily defined as an 'ongoing learning process'. In the development of our approach we focused on theories on learning and on the effectiveness of educational interventions. Consistent with literature, we did not find straightforward effectiveness of the educational approach.^{17 18 33} This may however not only be explained by the learning approach adopted, but also by problems related to the second half of our definition, namely 'meeting medical and societal standards in patient care'. The effectiveness of CPD depends on the commitment of the participants. Thus, the content of interventions and the accompanying willingness of participants to reach these standards is an important determinant in educational studies. Our study did not aim to determine the exact role of the intervention's content, and more determinants may have influenced the effectiveness. We will however further reflect on this hypothesis as it may reveal important aspects in the context of CPD.

The major conclusion of these reflections is that, to enhance the effectiveness of CPD, we should in the current situation not primarily focus on educational methods, but mainly on aspects in the professional culture and organisation, both individually and collectively. We will consider three aspects that are related to the standards that CPD aims at and may interfere with the effectiveness of educational support. These are: the definition of standards for desired care, the acceptance of these standards by individual professionals and the motivation to reach these standards.

The first problem in the effectiveness of CPD may be that standards of desired performance are often not described sufficiently. In general practice, the broad domain has just recently been defined more precise.^{34 35} On specific subdomains, guidelines are available, but there is no agreement to what extent adherence to these recommendations is desirable. For an individual GP, it is difficult or maybe even impossible to fully cover the whole domain of general practice, and in the current situation individual choices are made which aspects out of the domain are dealt with and to what extent.

This is frustrating for GPs with high internal standards, and comfortable for those who are not over-ambitious. In the current situation, self-directed learning activities lead to a divergence in the performance of individual health care professionals. For effective CPD, we need to formulate standards that are at least specific, realistic and acceptable for GPs. Our study towards a criterion-referenced standard for doctor-patient communication is an example that proves that this is possible under the right conditions.

A second problem in the effectiveness of CPD may be that existing standards on desired performance are free of obligation. Effective education requires of GPs to collectively define their main tasks and the level of performance that can be expected, and of individual GPs to conform to these standards.

In the current professional culture however, GPs have a considerable individual professional autonomy to define their own standards. Autonomy is a privilege of professionals, necessary to perform a complex task.^{14 36-39} Professional autonomy however is misunderstood, if it permits demonstrated deficiencies in performances to be neglected because of individual preferences or beliefs, as was the case in ENT-disorders, an undisputable element of general practice. Individual professional autonomy has wrongfully led to a situation in which GPs, with different conceptions of priorities in their tasks, are dispersed by self-directed learning activities over the broad domain of general practice, with the logical consequence of a large variation in actual performance. This situation is not compatible with the desire to achieve certain standards in the quality of care. We therefore require a new type of professionalism, in which the individual professionals allow themselves to be 'directed' to developing competencies and performances that have been agreed on collectively.

Also, GPs' motivation to reach existing standards may be insufficient. Current improvement activities of individual GPs and practices are mainly driven by internal motivation, which appears to be unsatisfactory. In the study, participants were difficult to be motivated to change aspects of their performance in ENT-disorders, despite the low scores on the assessment. No external motivating factors to improve the quality of care

are present here; whether for example prescriptions of antibiotics are in conformity with guidelines or not, remains invisible for outsiders. If differences can be made visible, external motivation in the form of social influence of colleagues, legislation or financial incentives is possible.

A number of developments in the direction of what has been discussed above are already taking place, e.g. the development of indicators, a more specific description of the domain of general practice and the possibility of financial incentives for specific performances.³⁵ The challenge will be to change aspects in the professional culture thoughtfully and in the right order. Therefore it is necessary to have insight in the myriad of factors that determine the effectiveness of CPD.

Doctors' learning and improving takes place in a complex, multifaceted reality and many factors influence the effectiveness of CPD-interventions.⁴⁰ A concept towards learning of professionals has to cover these factors. The adopted learning approach is only one of them (see figure). Other factors already discussed above, are related to the learners, the learners' professional contexts, the learning contents and undoubtedly several more. It is likely that there is also a considerable interaction between these factors, as can be easily illustrated.¹⁶ For example, motivation for a specific topic will influence the effectiveness of learning activities in a positive way, and vice versa: effective learning will have a positive effect on the motivation to learn and improve and use methods that have proven to be useful.

Further development of the effectiveness of professional learning requires a model that comprises these aspects in the context, the content and the educational approach as well as their mutual interactions. For future consideration, we propose a model for discussion that contains a number of these factors and their possible interactions (see figure).

This model acknowledges the multifaceted reality of professional learning. Central in this figure is the effectiveness of CPD. The learning method is just one of the factors that determine the effectiveness of professional learning. A number of other factors are also given. Motivation to learn is put as a key factor in the model, and many other aspects influence it.

The lower half of the figure comprises aspects in the effectiveness of professional learning that were integrated in our intervention. On the right side of the model, ideal and actual practice are compared to identify learning needs. The available room for improvement will directly determine the effectiveness of CPD-activities. The identified learning needs may also affect the motivation of GPs: a small desired improvement may not be inspiring, and if the analysed gap is perceived as impossible to tackle, professionals may feel discouraged.^{41 42} In the small-group meetings, GPs were encouraged to analyse barriers and facilitators to change. These barriers and facilitators directly influence the effectiveness of improvement activities, and the motivation to change.

Figure. Concept of a model with factors influencing the effectiveness of educational support.

In the upper half of the model, a number of aspects that influence the effectiveness of CPD are given, that were encountered explicitly or implicitly in the study. These are individual aspects, such as the motivation to learn and improve. Also external motivating factors (e.g. social influence of peers or financial incentives) may influence the motivation to learn. We discuss some aspects in more detail.

Costs and benefits are the result of a trade-off between the perceived gains of achieving certain standards, for example in terms of patient satisfaction, and the required investment to reach a standard, for example the time needed or the possible barriers encountered. Participating GPs consider the costs and benefits of learning activities in an early stage and quickly decide whether they regard it worthwhile to invest time and efforts.⁴³ Set in the margin of the model, the professional culture is yet an important modifying factor that may influence the effectiveness of learning in several ways. Professional culture is defined by what doctors think, how they think and how they work together. It is no static feature, but a dynamic one and an evolving socially constructed reality that exists in the minds of health professionals, and that can be regarded as the 'normative' glue that allows group members to communicate and work together effectively.⁴⁴ One

important aspect of the professional culture with regard to the effectiveness of future CPD will be how collective and individual educational tasks and responsibilities are to be defined. These have to be translated in practical terms like accreditation or certification of doctors or practices.

To improve professional learning, we cannot put all our eggs in one basket, but we have to consider numerous factors and use several methods.⁸ Collectively, the professional group should, besides further development of educational methods, look critically at the current culture of learning, at the definition of standards of desired care and at factors that motivate professionals to reach these standards.^{14 37 45} Individually, health care professionals should further improve their skills to learn and improve and to develop their individual stance towards collectively defined standards.

Implications for research on CPD

More research is needed to further improve educational interventions. This encompasses new assessments, the development of standards of adequate performance as well as learning approaches that fit in current routines and are attractive for GPs to participate in. In order to be able to evaluate the effectiveness of new and probably complex interventions, the most crucial and disputable elements must be clearly defined. These should be studied with small contrast between study arms. In our study, a number of factors may explain differences, as the topic, the assessments and the effect measures are nested within study arms. An interesting - but also expensive - approach would be a four-arm study, focusing on doctor-patient communication and ENT-disorders, that are both in one half assessed on the level of performance, and in the other half on the level of knowledge. In future research, after having defined the crucial elements of multifaceted approaches, these elements should be studied in appropriate designs, such as factorial designs.

But most of all, studies are needed to understand the factors that influence the effectiveness of professional learning and their mutual interactions. Current research is often aimed at effectiveness of specific methods, or at finding the best approaches for specific clinical problems. But it proves to be difficult to define generic insights that support those involved in organisation of support of professional learning on the basis of the outcomes of these specific studies. Future research demands the development of a theoretical model that comprises relevant aspects as discussed above, and which is currently not available. A model should contain the most important theoretical constructs and their mutual interactions. These concepts should be based on theories that acknowledge the complexity of learning and behaviour change and the importance of values and motivation.^{46 47} Recently, an interesting first step towards such a model has been made by defining construct out of psychological theory.⁴⁸ Examples of important constructs and factor that need further study are professionalism and professional culture, as the current literature on these themes is mainly based on opinions, and not on evidence. What professional characteristics are needed in current health care for optimal performance?

What is the current professional attitude towards learning? How can these be changed? What are side effects of interventions on professionalism itself and on motivation?⁴⁹ The results from these studies will provide the information needed to develop educational interventions and an educational context that fit in with the medical and societal demands with regard to effective and efficient health care.

LITERATURE

1. Grant J. Learning needs assessment: assessing the need. *BMJ* 2002;324:156-9.
2. Norman GR, Shannon SI, Marrin ML. The need for needs assessment in continuing medical education. *BMJ* 2004;328:999-1001.
3. Borgiel AE, Williams JI, Davis DA, Dunn EV, Hobbs N, Hutchison B, et al. Evaluating the effectiveness of 2 educational interventions in family practice. *CMAJ* 1999;161:965-70.
4. Ram PM. Comprehensive Assessment of General Practitioners. Universiteit Maastricht, 1998.
5. Hobma S, Ram P, van der Vleuten C, Grol R. Gebruiken huisartsen toetsen voor kwaliteitsverbetering? *Huisarts & Wetenschap* 2002;45:726-9.
6. Batalden PB, Stoltz PK. A framework for the continual improvement of health care: building and applying professional and improvement knowledge to test changes in daily work. *Jt Comm J Qual Improv* 1993;19:424-47.
7. Bateman H, Kinmonth AL. Journeys and pathways: exploring the role of professional development advice and educational guidance for practitioners expressing interest in research. *Med Educ* 2001;35:49-55.
8. Grol R. Improving the Quality of Medical Care: building bridges among professional pride, payer profit and patient satisfaction. *JAMA* 2001;286:2578-85.
9. Birkmeyer JD, Siewers AE, Finlayson EVA, Stukel TA, Lucas FL, Batista I, et al. Hospital volume and surgical mortality in the United States. *N Engl J Med* 2002;34:1128-39.
10. Birkmeyer JD, Stukel TA, Siewers AE, Goodney PP, Wennberg DE, Lucas FL. Surgeon volume and operative mortality in the United States. *N Engl J Med* 2003;349:2117-27.
11. Lerut T. The surgeon as a prognostic factor. *Ann Surg* 2000;232:729-32.
12. Carter DC. The surgeon as a risk factor. *BMJ* 2003;326:832-3.
13. Albanese MA. Developing effective, efficient and practical methods of assessing the professional skills of physicians in practice. *Med Educ* 2004;38:4-5.
14. Wynia MK, Latham SR, Kao AC, Berg JW, Emanuel LL. Medical Professionalism in Society. *N Engl J Med* 1999;341:1612-6.
15. Albanese MA. Treading tactfully on tutor turf: does PBL tutor content make a difference? *Med Educ* 2004;38:916-20.
16. Van Berkel HJM, Schmidt HG. Motivation to commit oneself as a determinant of achievement in problem-based learning. *Higher Education* 2000;40:231-42.
17. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing Physician Performance. *JAMA* 1995;274:700-705.
18. Oxman AD, Thomson MA, Davis DA, Haynes RB. No magic bullets: a systematic review of 102 trials of interventions to improve professional practice. *CMAJ* 1995;153:1423-1431.
19. Wensing M, Grol R. Single and combined strategies for implementing changes in primary care: a literature review. *Int J Qual Health Care* 1994;6:115-132.
20. Murray TS, Campbell LM. Finance, not learning needs, makes general practitioners attend courses: a database survey. *BMJ* 1997;315:353.
21. Pitts J, Vincent S. General practitioners' reasons for not attending a higher professional education course. *Br J Gen Pract* 1994;44:271-3.
22. Grol R, Wensing M, Eccles M. *Improving patient care?* Oxford: Elsevier, 2005.
23. Wensing M, Hulscher M, Laurant M, Grol R. Identifying Barriers and Facilitators for Changing Professional Practice. Nijmegen: WOK, 1999.
24. Thorsen T, Mäkelä M. *Changing Professional Practice. Theory and Practice of Clinical Guidelines Implementation*. Copenhagen: Danish Institute for Health Services Research and Development, 1999.

25. Flottorp S, Havelsrud K, Oxman AD. Process evaluation of a cluster randomized trial of tailored interventions to implement guidelines in primary care--why is it so hard to change practice? *Fam Pract* 2003;20:333-9.
26. Flottorp S, Oxman AD, Havelsrud K, Treweek S, Herrin J. Cluster randomised controlled trial of tailored interventions to improve the management of urinary tract infections in women and sore throat. *BMJ* 2002;325:367.
27. Rethans J-J. Does Competence predict Performance? Universiteit Maastricht, 1991.
28. Ram P, Vleuten Cvd, Rethans J, Schouten B, Hobma S, Grol R. Assessment in general practice: the predictive value of written-knowledge tests and a multiple-station examination for actual medical performance in daily practice. *Med Educ* 1999;33:197-203.
29. Ram P, Van der Vleuten C, Rethans JJ, Grol R, Aretz K. Assessment of practicing family physicians: comparison of observation in a multiple-station examination using standardized patients with observation of consultations in daily practice. *Acad Med* 1999;74:62-69.
30. Sargeant JM, Mann KV, Ferrier SN, Langille DB, Muirhead PD, Hayes VM, et al. Responses of rural family physicians and their colleague and coworker raters to a multi-source feedback process: a pilot study. *Acad Med* 2003;78:S42-4.
31. Davis W, Nairn R, Paine ME, Anderson RM, Oh MS. Effects of expert and non-expert facilitators on the small-group process and on student performance. *Acad Med* 1992;67:470-4.
32. Dijkstra R. Implementing diabetes guidelines at outpatient clinics. Patient-centred and professional-directed approaches. UMC St Radboud, 2004.
33. Grimshaw JM, Thomas RE, MacLennan G, Fraser C, Ramsay CR, Vale L, et al. Effectiveness and efficiency of guideline dissemination and implementation strategies. *Health technology assessment Winchester, England* 2004;8:iii-iv, 1-72.
34. *Toekomstvisie huisartsenzorg. Huisartsenzorg in 2012: medische zorg in de buurt.* Utrecht: LHV/NHG, 2002.
35. *Het aanbod aan huisartsgeneeskundige zorg 2004.* Utrecht: LHV, 2004.
36. Irvine D. The performance of doctors. I: Professionalism and self regulation in a changing world. *BMJ* 1997;314:1540-2.
37. Irvine D. The performance of doctors: the new professionalism. *Lancet* 1999;353:1174-7.
38. Casalino LP. The unintended consequences of measuring quality on the quality of medical care. *N Engl J Med* 1999;341:1147-50.
39. Rothman DJ. Medical professionalism - focusing on the real issues. *N Engl J Med* 2000;342:1284-6.
40. Smits PB, Verbeek JH, Nauta MC, Ten Cate TJ, Metz JC, van Dijk FJ. Factors predictive of successful learning in postgraduate medical education. *Med Educ* 2004;38:758-66.
41. Fox RD, Mazmanian PE, Putnam RW. *Changing and learning in the lives of physicians.* New York: Praeger, 1989.
42. Handfield-Jones RS, Mann KV, Challis ME, Hobma SO, Klass DJ, McManus IC, et al. Linking assessment to learning: a new route to quality assurance in medical practice. *Med Educ* 2002;36:949-58.
43. Slotnick HB. How doctors learn: physicians' self-directed learning episodes. *Acad Med* 1999;74:1106-17.
44. Hudelson PM. Culture and quality: an anthropological perspective. *Int J Qual Health Care* 2004;16:345-6.
45. Irvine D. Doctors in the UK: their new professionalism and its regulatory framework. *Lancet* 2001;358:1807-10.
46. Covington MV. Goal theory, motivation and school achievement: an integrative view. *Annu rev. Psychol.* 2000;51:171-200.
47. Eccles JS, Wigfield A. Motivational beliefs, values, and goals. *Annual Review of Psychology* 2002;53:109-32.
48. Michie S, Johnston M, Abraham C, Lawton R, Parker D, Walker A, et al. Making psychological theory useful for implementing evidence based practice: a consensus approach. *Qual Saf Health Care* 2005;14:26-33.
49. Marshall M, Harrison S. It's about more than money: financial incentives and internal motivation. *Qual Saf Health Care* 2005;14:3-4.

 

SUMMARY

Introduction

This thesis describes a study on continuing professional development (CPD) of general practitioners in The Netherlands. CPD is the necessary ongoing education of doctors during their careers in order to maintain the competence to practise and to optimise patient care. It has become clear that doctors need support in their continuing professional development, and this has been the focus of extensive educational research and development during the last decades. None of the developed theories, approaches or instruments has yet demonstrated to come forward to the expectations with regard to their effectiveness. The aim of this study is to further develop and ground an approach to support general practitioners (GPs) in their continuing professional development on the basis of the available insights, and to investigate the effectiveness and feasibility of this approach. The specific research questions were: (1) What are the features of an approach that can be expected to effectively and efficiently support CPD of medical professionals; (2) is this approach more effective in changing GPs performances towards desired standards than current educational approaches; (3) is this approach feasible, i.e. what are the needs in terms of time and money and do GPs appreciate the developed approach?

Chapter 2

This chapter describes the development of a multifaceted intervention to support continuing professional development of Dutch GPs. The development was based on literature, on needs and barriers as encountered by GPs and on the opinions of experts involved in educating medical professionals.

Starting point for the development are the concepts and the underlying theories as described in literature on CPD. Combining these theories with the evidence on the effectiveness of educational methods, we propose directed self-learning as a concept that depends on both self-initiated activities, but that also supports individual GPs in specific aspects of CPD.

The intervention aims to be more effective than current approaches by tailoring learning activities to objective individual learning needs of GPs. The first step in our model is a needs assessment to support GPs in identifying what aspects of performance need improvement. On the basis of the assessment results and predefined criterion-referenced standards of adequate performance, a subdomain for further improvement is selected from the assessed domain. GPs receive feedback on their results and the selected subdomain. Consequently, they receive small-group education, supervised by a tutor who is an expert on the topic of interest.

In a small-scale pilot, this approach appeared to be acceptable and feasible for participating GPs, but some adjustments were necessary to enable the intervention to be used on a larger scale. The main change is to support GPs in the process of quality improvement only, but to consider the contents and targets of their activities as their own professional responsibility.

Chapter 3

This chapter addresses GPs' use of available assessment methods, GPs use of feedback for CPD, and the needs and barriers they encounter. This information was necessary to tailor the intervention to our specific situation of general practice in The Netherlands. Assessment is regarded as a key element of CPD, but little is known about its current use and educational impact. We studied two samples of Dutch GPs. The actual use of assessment methods and the barriers encountered in their use were investigated by semi-structured interviews by telephone with a random group of GPs. The use of feedback on assessment results for CPD was investigated with a written questionnaire, sent to GPs with extensive experience with assessment.

It appears that assessments are hardly used for professional learning. More than half of the randomly selected GPs had little or no experience with assessment. Of the GPs who used assessments, more than 80% had not used the feedback for CPD. The limited use is due to problems in using the assessment methods and difficulties in using the results. GPs have a positive attitude to assessment, but want better feedback and need support. Workload reduction, improvement of practice organisation, enhancing the support of GPs in using the methods, and legislation are seen as possible solutions by the GPs, to enhance the use of assessments.

Chapter 4

This chapter deals with the definition of a criterion-referenced standard, a requirement for an assessment-based CPD approach that is described in chapter 2. Criterion-referenced standards for assessing performance in real practice of GPs should be available to identify learning needs for CPD. The applicability of common standard-setting procedures in authentic assessment however has not been investigated.

To set a standard for assessment of GP-patient communication with video observation of daily practice, two well-known examples of two different standard setting approaches were investigated; an Angoff-procedure was applied to written cases, and a borderline regression method was applied to a sample of videotaped consultations of GPs. The procedures and outcomes were evaluated by the applicability of the procedure, the reliability of the standards and the credibility as perceived by assessed GPs.

Both methods appeared to be applicable and reliable; the obtained standards are credible according to the GPs. The context in which the standard is to be used is important because methods differ in practical aspects.

Chapter 5

This chapter describes the effectiveness of the developed intervention in improving doctor-patient communication. This is an essential component of general practice and improvement of GPs' communication patterns is an important target of training programmes. Available studies have so far failed to provide conclusive evidence on the effectiveness of educational interventions to improve doctor-patient communication. To examine the effectiveness of the CPD-approach a RCT was performed. In the interven-

tion group, observing authentic consultations in their own surgery identified learning needs. This performance assessment was followed by personal feedback and structured activities in small-group meetings, aimed at remedying the identified shortcomings. The control group received existing educational materials.

The scores in the intervention group demonstrated a significant improvement compared with those of the control group. The effect size was moderate to large. The level of participation significantly contributed to the effectiveness. Most improvement was seen on patient centred communication skills.

The conclusion is that the approach of structured individual improvement activities based on performance assessment is more effective in improving communication skills than current educational activities.

Chapter 6

This chapter describes the effectiveness of the intervention to improve the care for ENT-disorders. The effectiveness was studied using indicators on structure and process, derived from general practice guidelines. These were sampled by questionnaires on practice organisation and equipment, by GPs' prospective recordings of performance and by collecting prescription data. Participants in the intervention group were allocated to the small-group meetings on the basis of a written knowledge test. Participants in the control arm received existing educational materials on ENT-disorders.

A small but non-significant improvement on primary effect measures was found comparing the intervention group with the control group. Subgroup analyses showed no effect of the invested time in the intervention.

The conclusion is that the intervention did not demonstrate effectiveness in improving relevant aspects of care for ENT-disorders. This can only partly be explained by shortcomings in our study, such as insufficient power or insensitivity to the effect measures or by a lack of face-validity of the knowledge test used. A probable important factor influencing effectiveness was the topic (ENT-disorders) itself, for which GPs were not highly motivated.

Chapter 7

This chapter focuses on the appreciation and feasibility of the educational intervention. The intervention studied is an example of the growing tendency to develop more complex interventions for CPD in order to enhance effectiveness. Besides effectiveness, feasibility and the appreciation of stakeholders are key features for implementation in daily educational routines. The ability of GPs and supporting staff to perform the intervention, the costs per hour and the participants' appreciation of (aspects of) the educational intervention were studied.

It was found that GPs are able to perform a CPD-intervention that starts with a needs assessment and that subsequently supports the individual self-directed learning process. Participants need on average 22.3 hours for the assessments, small-group meetings and work in their practices. The costs are € 117.56 per hour. The mean appreciation is 6.8 on

a 10-point scale. Appreciation of and participation in the intervention appear to be dependent on the topic studied.

It is concluded that the approach towards continuing professional development is feasible and acceptable. A context is required in which sufficient resources are available with respect to budget, educational materials and skilled support staff. Furthermore, GPs must be interested in the topics studied and probably also in the specific approach.

General discussion

In the general discussion, the conclusion is that, taken the findings together, the study does not provide evidence that the intervention is an approach that is superior to current educational routines. The effectiveness was ambiguous; the appreciation of stakeholders moderate and the needs in terms of time and costs were considerable. We discuss whether the results regarding the effectiveness must be contributed to the underlying concept, to details in the intervention or to flaws in the study design.

The concept contains valuable and necessary elements such as support in needs assessment needs and improving care. It does, however not fit in the current professional culture towards learning. The concept was elaborated into an intervention, and pragmatic choices had to be made here. Aspects in the intervention, for example the assessments used or the tutor-expertise, may have influenced the results. Also shortcomings in the study design and effect measures and their possible effects are discussed. The importance of the discussed aspects in the intervention or study design must however not be overestimated and do not alter the conclusions about the concept.

For future CPD this implies that to support professional learning, interventions should aim at effective and feasible educational approaches, but also takes the professional context and culture into account. CPD needs further development of methods, such as performance assessments, feasible feedback and more 'improvement knowledge' in the target group. Aspects in the current professional culture that need to be studied and altered are: standards of desired care must clearly defined, the concept of individual and collective professional autonomy must be reconsidered and more external motivating stimuli to improve are necessary.

Inleiding

Dit proefschrift beschrijft een onderzoek naar 'continuing professional development' - letterlijk vertaald: voortdurende professionele ontwikkeling - van Nederlandse huisartsen. Artsen moeten zich tijdens hun werkzame leven blijven scholen en ontwikkelen om goede patiëntenzorg te leveren. Hiervoor hebben zij ondersteuning nodig. De laatste decennia is veel onderzoek gedaan hoe deze ondersteuning het beste kan worden vormgegeven. Geen van de ontwikkelde methodes voldoet echter aan de verwachtingen. In deze studie is een benadering ontwikkeld die artsen ondersteunt in hun professionele ontwikkeling, en is de effectiviteit en haalbaarheid hiervan onderzocht. De onderzoeksvragen waren: (1) Hoe ziet een benadering eruit waarvan verwacht kan worden dat deze effectief en efficiënt de continue professionele ontwikkeling van (huis)artsen ondersteunt; (2) is deze benadering effectiever dan bestaande nascholingsmethoden; (3) is deze benadering haalbaar: wat is de benodigde investering in tijd, geld, mensen en middelen, en hoe waarderen huisartsen de nieuwe benadering?

Hoofdstuk 2

Dit hoofdstuk beschrijft de ontwikkeling van een interventie om Nederlandse huisartsen bij hun professionele ontwikkeling te ondersteunen. De ontwikkeling van de interventie is gebaseerd op de literatuur over professionele ontwikkeling, op een studie naar de behoeften en knelpunten die door huisartsen worden gerapporteerd (beschreven in hoofdstuk 3) en op de meningen van deskundigen, betrokken bij de nascholing van artsen. Uitgangspunt voor de ontwikkeling van de interventie waren de onderwijskundige theorieën waarnaar in de literatuur over professionele ontwikkeling verwezen wordt. Daarnaast werd gebruik gemaakt van inzichten uit onderzoek naar de effectiviteit van onderwijskundige interventies. Op basis van theorieën en onderzoeksgegevens werd de hypothese ontwikkeld dat professioneel leren moet worden vormgegeven als 'gestuurde zelfeducatie'. In dit concept is de deelnemend arts zelf verantwoordelijk voor een belangrijk deel van de leeractiviteiten, maar wordt ook ondersteund in een aantal aspecten van het leren.

De interventie brengt de sterke en zwakke punten van individuele huisartsen in kaart, en ondersteunt hen vervolgens in de verbetering van hun dagelijks handelen. De eerste stap bestaat uit toetsing van de objectieve leerbehoeften. De toetsresultaten worden vergeleken met tevoren vastgestelde normen voor adequate scores en zo worden onderwerpen geselecteerd voor verdere activiteiten. Huisartsen krijgen feedback over hun toetsresultaten en de geselecteerde onderwerpen. Vervolgens komen zij een aantal malen bijeen in kleine groepen onder leiding van een inhoudsdeskundige tutor om te werken aan verbetering van de kwaliteit van hun dagelijks handelen.

In een pilotstudie bleek de opzet haalbaar en acceptabel voor huisartsen. De structuur van de groepsbijeenkomsten werd naar aanleiding van de pilotstudie aangepast om de haalbaarheid op grotere schaal te verbeteren: huisartsen werden nog wel ondersteund in het proces van kwaliteitsverbetering, maar de vakinhoudelijke invulling van de bijeenkomsten werd de verantwoordelijkheid van de deelnemers zelf.

Hoofdstuk 3

Dit hoofdstuk gaat over het gebruik van toetsen door huisartsen, en de behoeften en knelpunten die zij hierbij ondervinden; deze gegevens waren nodig bij de ontwikkeling van de interventie, beschreven in hoofdstuk 2. Toetsing wordt beschouwd als een belangrijk onderdeel van professioneel leren, maar er is weinig bekend over het daadwerkelijk gebruik van toetsen en toetsresultaten.

Er werden twee geselecteerde groepen Nederlandse huisartsen onderzocht. Het gebruik van de beschikbare toetsen werd onderzocht door telefonische semi-gestructureerde interviews met een aselect gekozen groep Nederlandse huisartsen. Het gebruik van toetsresultaten werd onderzocht bij huisartsen die uitgebreide ervaring hadden met meerdere toetsmethoden met een schriftelijke vragenlijst.

Meer dan de helft van de geïnterviewde huisartsen had geen of zeer beperkte ervaring met toetsen. Van de groep huisartsen die uitgebreide ervaring had met toetsing rapporteerde meer dan 80% de toetsresultaten niet te hebben gebruikt voor verdere leeractiviteiten.

Toetsen worden dus nauwelijks gebruikt voor professionele ontwikkeling. Dit is deels te wijten aan de toetsmethoden zelf, en deels aan problemen in het gebruik van de resultaten voor verdere leeractiviteiten. Huisartsen staan positief tegenover toetsing, maar willen betere feedback en begeleiding bij het gebruik van toetsresultaten. Om dit te ondersteunen noemen huisartsen de volgende oplossingen: vermindering van de werkdruk, verbetering van de praktijkorganisatie, meer ondersteuning bij leeractiviteiten en regelgeving.

Hoofdstuk 4

Dit hoofdstuk beschrijft de ontwikkeling van een absolute norm voor de toetsing van arts-patiënt communicatie door video-observatie in de dagelijkse praktijk. In hoofdstuk 2 is beschreven dat deze normen nodig zijn om de toetsresultaten te kunnen interpreteren. Er zijn een aantal normstellingprocedures in de literatuur beschreven en onderzocht, maar de bruikbaarheid van deze methoden voor toetsing van het dagelijks handelen is niet bekend.

Om een norm voor toetsing van het communicatief handelen van huisartsen te bepalen werden twee bekende representanten van verschillende normstellingbenaderingen onderzocht; een Angoff-procedure werd toegepast op geschreven casus, en een borderline-regressie methode werd toegepast op video-opnamen uit de dagelijkse praktijk van huisartsen. De verschillende benaderingen werden beoordeeld aan de hand van de toepasbaarheid en de betrouwbaarheid van de procedures, en de geloofwaardigheid van de norm in de ogen van getoetste huisartsen.

Beide methodes bleken toepasbaar en betrouwbaar, en beide normen waren voor huisartsen geloofwaardig. Voor de keuze van een methode is vooral het doel van de standaard belangrijk, omdat de methoden verschillen in een aantal praktische aspecten.

Hoofdstuk 5

Dit hoofdstuk gaat in op het effect van de ontwikkelde interventie op de communicatie tussen arts en patiënt. Dit is een belangrijk onderdeel van het huisartsgeneeskundig domein. Beschikbare studies laten zien dat scholing van de communicatieve gewoonten van praktiserende artsen vaak geen of weinig effect heeft. De effectiviteit werd onderzocht in een gerandomiseerde studie. De artsen in de interventie- en controlegroep werden aan het begin en einde van de studie met behulp van video geobserveerd in de eigen praktijk. De interventiegroep kreeg feedback op deze observaties, waarbij de scores werden gerelateerd aan de norm die was verkregen met de Angoff-procedure, zoals beschreven in hoofdstuk 4. De laagscorende huisartsen kwamen in kleine groepen bij elkaar om te werken aan verbetering van hun dagelijks handelen. De controlegroep ontving bestaande educatieve materialen van het Nederlands Huisartsen Genootschap, en kreeg pas na afloop van de studie feedback.

De scores van de deelnemers in de interventiegroep lieten een significante verbetering zien ten opzichte van de scores van de deelnemers in de controlegroep. Het effect van de interventie was matig tot groot. De mate waarin huisartsen deelnamen aan de bijeenkomsten droeg bij aan het effect. De effecten werden met name gezien op de door de deelnemers gekozen patiëntgerichte items.

De conclusie is dat een benadering die eerst zwakke plekken in het dagelijks handelen identificeert en vervolgens stapsgewijs aanpakt, effectiever is in het verbeteren van arts-patient communicatie dan beschikbare methoden.

Hoofdstuk 6

Dit hoofdstuk beschrijft het effect van de ontwikkelde interventie op het gebied van keel-, neus- en ooraandoeningen (KNO-aandoeningen). De effectiviteit werd gemeten door aan de start en het einde van de studie indicatoren, gebaseerd op de NHG-Standaarden, te bepalen die structuur en proces van de zorg op het gebied van KNO-aandoeningen in kaart brengen. Informatie over praktijkvoering werd verkregen met een schriftelijke vragenlijst. De adherentie aan NHG-Standaarden werd onderzocht met zelf-registratielijsten die door deelnemers werden ingevuld na consulten waarin KNO-problemen aan de orde kwamen. Tenslotte werden prescriptiecijfers over voor KNO-aandoeningen relevante medicatie verzameld. In de interventiegroep werden deelnemers geselecteerd met behulp van een schriftelijke kennistoets. De controlegroep kreeg bestaande materialen van het Nederlands Huisartsen Genootschap.

Er werden geen significante verschillen in de primaire effectmaten gevonden tussen de interventie en de controlegroep. De resultaten van de kennistoets hadden tot gevolg dat een groot deel van de deelnemers in de interventiegroep werden toegewezen aan kleine groepen die zich richtten op verbetering van de zorg op KNO-gebied, maar analyse liet geen effect van deze bijeenkomsten zien.

De conclusie is dat geen effectiviteit kon worden aangetoond van de ontwikkelde interventie op het gebied van KNO-aandoeningen. Dit kan mogelijk deels verklaard worden door tekortkomingen in de interventie, zoals de gebruikte kennistoets en beperkingen in

de effectmaten. Een belangrijke factor in het ontbreken van effectiviteit was waarschijnlijk het onderwerp, KNO-aandoeningen, zelf, omdat huisartsen matig gemotiveerd waren tot verbetering.

Hoofdstuk 7

In dit hoofdstuk wordt de waardering en de haalbaarheid van de interventie beschreven. De ontwikkelde interventie is een voorbeeld van de trend om steeds omvangrijker en ingewikkelder educatieve benaderingen toe te passen om de effectiviteit te verbeteren. Maar naast effectiviteit zijn haalbaarheid en de waardering door de doelgroep belangrijk voor de implementatie van onderwijskundige benaderingen in de nascholing van artsen. Daarom werden uitvoerbaarheid, benodigdheden, kosten en waardering door deelnemers onderzocht, zowel voor de interventie in zijn geheel als voor de verschillende onderdelen ervan.

Huisartsen zijn in staat de interventie uit te voeren. In totaal kost de interventie de deelnemers gemiddeld 22.3 uur. De kosten per uur zijn € 117.56. De gemiddelde waardering is een 6.8 op een 10-punts schaal. De waardering en de participatie lijken afhankelijk te zijn van het onderwerp dat wordt behandeld.

De conclusie is dat onze benadering haalbaar en acceptabel is. Implementatie van de benadering vereist een structuur waarin voldoende geld, materialen (zoals toetsen) en getrainde ondersteuning (zoals tutores) beschikbaar zijn. Ook moeten huisartsen geïnteresseerd zijn in de behandelde onderwerpen, en waarschijnlijk ook in de benadering zelf.

Discussie

Alles bij elkaar genomen is de conclusie dat er geen bewijs is geleverd dat de ontwikkelde interventie beter is dan op dit moment beschikbare methoden. De interventie laat een wisselende effectiviteit zien, de waardering door huisartsen is redelijk en de benodigdheden zijn aanzienlijk. Dit is reden om het onderliggende concept, de uitwerking van het concept in een interventie, en de studieopzet kritisch te beschouwen.

Het concept bevat een aantal waardevolle en noodzakelijke elementen, zoals het in kaart brengen van leerbehoeften en ondersteuning bij het verbeteren van het dagelijks handelen. Het concept lijkt echter onvoldoende te passen in de huidige professionele cultuur van nascholing en kwaliteitsverbetering. De interventie, als uitwerking van het concept, was door een aantal noodzakelijke praktische keuzes mogelijk nog niet optimaal. Bijvoorbeeld het gebruik van kennistoetsen of de beperkte training van de tutores kunnen de effectiviteit negatief beïnvloeden hebben. Ook zijn beperkingen in de effectmaten aan te wijzen. Ondanks deze tekortkomingen lijkt de conclusie dat de interventie onvoldoende in de huidige cultuur past gerechtvaardigd.

Voor de ondersteuning van professionele ontwikkeling in de toekomst moet een betere samenhang worden gezocht tussen educatieve interventies en professionele cultuur. Er zal verder moeten worden gezocht naar effectieve en haalbare ondersteuning van professioneel leren. Daarvoor zijn nieuwe en betere educatieve instrumenten, zoals bij-

voorbeeld toetsen van het dagelijks handelen en betere feedback nodig. Huisartsen moeten meer weten over methoden voor kwaliteitsverbetering, en zich meer bewust worden van tekortkomingen in de huidige nascholing. Maar betere ondersteuning alleen lijkt niet genoeg, ook in de professionele cultuur en organisatie zullen veranderingen nodig zijn. De belangrijkste aspecten hiervan zijn: het formuleren van heldere normen voor adequate zorg, veranderingen in de professionele autonomie en met name in de verhouding tussen de individuele huisarts en de beroepsgroep, en beloning voor goede kwaliteit.

APPENDIX

MAAS-Global rating list for consultation skills of doctors addressing doctor-patient communication. Each item is scored on a scale ranging from 0 to 6. For item 2 and 4 the rating "not applicable" is an additional option. The complete checklist and criteria list are available at http://www.hag.unimaas.nl/Maas-global_2000.

Communication skills for each separate phase

1. Introduction

- giving the patient room to tell his story
- general orientation on the reason for visit
- asking about other reasons for visit

2. Follow-up consultation

- naming previous complaints, requests for help and management plan
- asking about adherence to management plan
- asking about the course of the complaint

3. Request for help

- naming requests for help, wishes or expectations
- naming reasons that prompted the patient to come now
- completing exploring request for help

4. Physical examination

- instructions to the patient
- explanation of what is being done
- treating the patient with care and respect

5. Diagnosis

- naming findings and diagnosis/hypothesis
- naming causes or the relation between findings and diagnosis
- naming prognosis or expected course
- asking for patient's response

6. Management

- shared decision-making, discussing alternatives, risks and benefits
- discussing feasibility and adherence
- determining who will do what and when
- asking for patient's response

7. Evaluation of consultation

- general question
- responding to requests for help
- perspective for the time being

General communication skills

8. Exploration

- exploring requests for help, wishes or expectations
- exploring patient's response to information given
- within patient's frame of reference
- responding to nonverbal behaviour and cues

9. Emotions

- asking about/ exploring feelings
- reflecting feelings (including nature and intensity)
- sufficiently throughout the entire consultation

10. Information giving

- announcing, categorizing
- in small quantities, concrete explanations
- understandable language
- asking whether the patient understands

11. Summarisations

- content is correct, complete
- concise, rephrased
- checking
- sufficiently throughout the entire consultation

12. Structuring

- logical sequence of phases
- balanced division of time
- announcing (history taking, examination, other phases)

13. Empathy

- concerned, inviting and sincerely empathetic in intonation, gesture and eye contact
- expressing empathy in brief verbal responses

AUTHORS

Drs. C.P.M. Aretz, research assistant, Department of General practice, Centre for Quality of Care Research, Maastricht University

Prof. dr. R.P.T.M. Grol, Professor of Quality in Health Care, Department of General practice, Centre for Quality of Care Research, Maastricht University

Prof. dr. G.G. van Merode, professor of Health Organisation Policy & Economics, Department of Health Organisation Policy & Economics, Maastricht University

Dr. H.G.A. Mokkink, methodologist, Department of General practice, Centre for Quality of Care Research, University of Nijmegen

Dr. ir. A.M.M. Muijtjens, research methodologist – statistician, Department of Educational Development & Research, Maastricht University

Dr. P.M. Ram, general practitioner – researcher, Department of General practice, Centre for Quality of Care Research, Maastricht University

Prof. dr. C.P.M. van der Vleuten, professor of education, Department of Educational Development & Research, Maastricht University

PUBLICATIONS

Chapter 2

Directed self-learning for medical professionals: development of an intervention to support CPD

S.O. Hobma, P.M. Ram, C.P.M. van der Vleuten, R.P.T.M. Grol

The pilot study described in this chapter was published as: Feasibility and acceptability of an integrated system of assessment and improvement on asthma/COPD and doctor-patient communication for general practitioners. *European Journal of General Practice* 2000 (September);6:18.

Chapter 3

GPs' use of assessments for continuing professional development

S.O. Hobma, P.M. Ram, C.P.M. van der Vleuten, R.P.T.M. Grol

Published as: Gebruiken huisartsen toetsen voor kwaliteitsverbetering? *Huisarts & Wetenschap* 2002;45 (December):726-9.

Chapter 4

Setting a standard for performance assessment of doctor-patient communication in general practice

S.O. Hobma, P.M. Ram, A.M.M. Muijtjens, R.P.T.M. Grol, C.P.M. van der Vleuten

Published in *Medical Education* 2004 (December);38:1244-52.

Chapter 5

Effectiveness in improving doctor-patient communication

S.O. Hobma, P.M. Ram, A.M.M. Muijtjens, C.P.M. van der Vleuten, R.P.T.M. Grol

Accepted for publication by the *British Journal of General Practice*

Chapter 6

Effectiveness in improving clinical performance on ENT-disorders

S. Hobma, P.M. Ram, C.P.M. van der Vleuten, H.G.A. Mookink, C.P.M. Aretz, R.P.T.M. Grol

Submitted

Chapter 7

Feasibility and appreciation

S.O. Hobma, P.M. Ram, G.G. van Merode, C.P.M. van der Vleuten, R.P.T.M. Grol

Published as: Feasibility, appreciation and costs of a tailored continuing professional development approach for general practitioners. *Quality in Primary Care* 2004 (December);12:271-8.

DANKWOORD

Dit proefschrift is de bekroning van enerverende, afwisselende, geconcentreerde en leerzame jaren, waar ik met grote tevredenheid op terugkijk. Dat is voor een belangrijk deel te danken aan de bijzondere, aardige en interessante mensen waarmee ik kon samenwerken. Die mensen wil ik hier bedanken.

Het begon en eindigt in de huisartsopleiding. Yvonne en Paul, misschien weten jullie het niet eens meer, maar de afsluitende gesprekken van de huisartsopleiding waren de opmaat voor dit onderzoek. Een gesprek met Yvonne, een telefoontje van Paul ('hé jongen, zoek je een baantje?'), een lange fax met het onderzoeksvoorstel naar de praktijk van Wim Ritzen (die nog altijd als een Torah-rol op mijn bureau ligt), en in enkele weken was de zaak geregeld. En nu, bij de afsluiting van het onderzoek heb ik een aanstelling bij dezelfde opleiding, en het voorrecht om ruimte te krijgen het proefschrift af te ronden. Dat was twee keer zeer belangrijk. Bedankt.

De begeleidingscommissie, Johannes Dalhuysen, Henk Mokkink, Richard Starmans, Arno Timmermans en Theo Voorn, leverde een onmisbare bijdrage met de enthousiaste discussies hoe de interventie en de studie moesten worden vormgegeven. Henk Mokkink hielp daarnaast bij het verwerken van de gegevens uit de zelfregistratie van huisartsen, en daar zijn vele dagen werk in gaan zitten. Henk bedankt daarvoor, zonder jouw ervaring en kennis waren we hier niet uitgekomen.

De globale opzet van de interventie werd uitgetest in de pilot. De deelnemende huisartsen Fons Alberts, Kick Hamers, Gaspard Knops, Kay Liedekerken, Ed Pustjens, Rob Pustjens, Rene van Hooren en Esmeralda Willems, de tutoren Henk Goettsch en Jean Muris en de observatoren Lou Zuijlen en Hanneke Hamers wil ik voor hun inzet van toen bedanken.

Het NHG speelde op verschillende manieren een belangrijke rol in het project. Ron Hellsloot en Ingeborg von Eugen wil ik bedanken voor de perfecte samenwerking rond de subsidieaanvragen en de tussenrapportages, altijd onder de tijdsdruk van de deadlines van ZonMW. Elly Bakker regelde de praktische benodigdheden, zoals de materialen voor de controlegroep. En Margriet Bouma en Saskia Mol zorgden ervoor dat we voor de kennistoetsen een degelijke absolute norm beschikbaar hadden. Dat was niet alleen veel werk, maar ook bijzonder spannend, omdat de norm van de KNO-toets volgens de planning pas in de week voordat we die nodig hadden beschikbaar zou zijn. Het lukte. Zonder deze samenwerking was het project niet denkbaar geweest.

Het geld was er, het plan, nu de mensen nog om mee te doen. Het creëren van nieuwe netwerken tussen mensen was een belangrijk deel van het project. Via de DistrictsHuisartsenVerenigingen kregen we toegang tot veel ervaring, enthousiasme en een bestaand netwerk dat daarvoor onmisbaar was. Pieter van Wijk, Emmy Derckx, Gerrit van Roekel en Henk Derks dachten mee over allerlei praktische aspecten en hielpen bij de werving. Daar gingen vele uren inzitten. Jullie hulp was zeer belangrijk. Er is voor jullie veel veranderd de laatste tijd, veel verloren ook denk ik, en ik vraag me af hoe onderzoekers met soortgelijke studies in de toekomst deze praktische zaken moeten oplossen. Zonder de deelnemers, van uiterst sceptisch ('ik neem alleen deel omdat ik niets van uw aannames geloof') tot zeer enthousiast ('zo moet het!'), was het project gestrand. Het aantal huisartsen dat geïnteresseerd was bleek aanzienlijk groter dan de beschikbare middelen. Ik wil de deelnemende huisartsen, P. Beelen, M. Bolsius-Franssen, M. Botden, M. Brueren, H. Bultman, M. Castendijk, P. Dalinghaus-Nienhuys, J. Dams, P. Dasen, C. de Kock, A. de Vocht, G. di Giulio, A. Gieles, P. Gruijters, W. Heres, O. Herijgers, J. Hoevenaars, M. Klomp, P. Meurs, M.J.L. Rodrigues de Miranda, J. Schelfhout, P. Schröder, H. Spiers, F. Titulaer, A. v.d Sande, H. van den Hoogen, J. v. Hoorn, R. Vergroesen, P. Verstijnen, E. Vogelzang, M. Becht, A. Bots, M. de Jonge, J. de Leur, S. de Ridder, S. de Swart, D. Engelen, J.H.A.H. Ferrée, M. Gerritse, J. Lips, S. Maas, A. Marinus, J. Oosterbaan, A. Rasing, M. Rubens, Th. Schure, R. Swank, K. Tersmette, M. Teunissen, P. van Bracht, L. van Cruchten, R. van Gestel, A. van Heijningen, C. van Jaarsveld, J. van Loon, C. v. Moorsel, W. van Osta, G. van Roekel, I. van Sluisveld, J. Versteeg, G. Spitz, J. v. Dongen, F. v.d. Ven, J. Wouters, B. Koopmans, G. Kusters, D. Snel, H. de Vries, T. Sledsens, M. Brassé, A. Storken, J. Reitsma-Maandag, J. Wijnhoven, C. Liedekerken, R. Pustjens, R. van Hooren, A. Alberts, C. Vincentie, H. van Rens, A. Ariëns, G. Wolfs, R. Hendriks, Th. v.d. Waart, C. Wijnands, R. Zonneveld, H. Ausems, G. Knops, J. Dellevoet, P. Schepers, F. Donders, N. Egelie, W. Nieuwdorp, B. Otten, V. Kaiser, J. Veldhuizen, T. van Dongen en M. van Dijk bedanken.

In de organisatie van het project nam de videotoets een prominente plaats in. De installateurs moesten, zo bleek, technische en diplomatieke vaardigheden bezitten. Ik dank Arn Bergevoet en Daan Matakana voor hun hulp bij het samenstellen van de videosets, en hun installaties in het eerste deel van het project. Ad Vorstenbosch en Erik Vervaart staken ziel en zaligheid in het project, en hun hulp was zeer belangrijk. Ook andere installateurs zoals E. Melman, M. Doorewaard, S. Harms, T. Harms, M. Huibers en diverse andere leverden goede bijdragen.

Daarna konden de observatoren aan het werk om de banden te scoren. Esther Tuerlings, H. Locht, Marjo van Bommel, P. de Wildt, C. van Gorp, Paul Wouda, Alfons Notermans, M. van Putten, Joke Koopmans, Jan van Rooij, Eveline van den Nieuwenhuijzen, Gerrit ten Oever, J. Naus, Antoon Visschedijk, Peter van Galen, Gerda Fijten, Lou Zuijlen, Renate Bongers, Henriette Swijgman, Leo Weusten, Rob Haesen, D. Mooren, G. Nijdam, Jef Swaans, M. van den Blink, R. Houppermans, Stefan Steins, G. Hardy, N. Jacobs-Jaspers, M. op den Kamp, George America, Marjan van de Mheen, H. Geurts van Kessel, Thijs Bakker, M. Smit, Joy Wong Chung, Peter van der Linden, Theo

Boogert, M. van den Goor, Frank van Horne, T. Smits, Maria Elbers en Annet Hoogveld hielpen hier mee.

Nadat de toetsen verwerkt waren gingen een aantal collega's op pad voor een ongewone klus: feedback geven. Esther Tuerlings, Jack van Son, Wendela van der Laan, Alfons Olde Loohuis, Ramon Aenderkerk en Phile Govaert bedank ik voor hun durf en hun inzet hierbij.

Vervolgens startten de groepsbijeenkomsten, in een roerige tijd. De meeste deelnemers hadden hun eerste bijeenkomst enkele uren nadat de Twin Towers in elkaar waren gezakt. Anthrax leek een reële dreiging, en pakketjes met banden en onderwijsmaterialen werden plots 'als verdacht aangemerkt', en bleven steken in opslagruimtes van het postbedrijf. Welke bias dit voor interventie- en controlegroep heeft opgeleverd zal nooit duidelijk worden. De tutores, Meggy van Kruijsdijk, Sylvia van Manen, Mark Nijboer, Leopold Messchaert, Arja van Wijgerden, Luc Maartens, Gerard Kummeling, Irene v.d. Heuvel, Alfons Notermans, Pieter Leijte, Antoon Visschedijk, Eline Becker, Peter Hezemans en Wim Verstappen brachten de aandacht van de deelnemers terug naar het project. Jullie inzet is door de deelnemers zeer gewaardeerd. Bedankt voor jullie hulp.

Bij het verzamelen en verwerken van de gegevens uit de hoofdstudie heb ik de noodzakelijke hulp gehad van een aantal mensen. Michel Wensing ondersteunde bij de powerberekening. Arno Muijtjens hielp, rustig, kritisch en nauwkeurig, met de analyses voor de normstellingsprocedures en het hoofdstuk over communicatie en schreef mee met de betreffende hoofdstukken. Arnold Kester gaf advies bij de analyses van het hoofdstuk over KNO-aandoeningen. Frits van Merode maakte mij vlot wegwijs in het beschrijven van de kostenaspecten en gaf snel en gedetailleerd commentaar op het hoofdstuk daarover. Alfons Schrooten wist uit de enorme hoeveelheid gegevens, die de aanschaf van nieuwe hardware noodzakelijk maakten, van de verzekeraars een werkbaar bestand te maken, en samen besteedden we veel tijd aan de formulering van de indicatoren. Paul Muijters leverde op diverse manieren bijdragen aan het onderzoek en dit proefschrift, zoals onder andere de indicatoren die we gezamenlijk gebruikten.

En dan de meer aardse zaken. Geld speelt geen rol voor een heer van stand, maar wel voor een onderzoeker. Toen onze plannen de beschikbare middelen van ZonMW dreigden te overstijgen stond het NHG financieel garant, zodat we het project toch konden uitvoeren zoals we dat graag wilden. Bij de verlenging van het project stond de huisartsopleiding daarvoor garant. Erie van den Heuvel en Ellen Breevoort ondersteunden ons bij de financiële aspecten en het contracteren van de vele medewerkers van het project. Joost Dormans hielp bij het maken van een planning voor het project. Het WOK-secretariaat, Karin Vaessen, Marie-Louise Dumont en Bernadette Zinzen, hielp met allerlei klussen. Gelukkig was daar ruimte voor, want de geschatte secretariale ondersteuning schoot schromelijk tekort. Aan het eind kreeg ik ook hulp van het secretariaat van de opleiding, Hub, Lilian, Alexandra, Mariet, Marleen en Helga. Collega's binnen en buiten de vakgroep gaven advies, commentaar of waren gewoon goed gezelschap. Paul Zwietering, als eerste en langstdurende kamergenoot: ik mis de discussies en het is nog steeds jammer dat, net toen ik meer rotzooi in de kamer had verzameld dan jou, we

van elkaar moesten scheiden. En natuurlijk Wim, Ben, Rogier, Mark, Ine (de hoeksteen van de vakgroep), Trudy, Caro, Marjan, Babette, Paul K., Paddy (v.L.), Sandra, Marcus, Edwin, Arno, Teun, Saskia, Maaïke, Johannes en anderen, bedankt voor zinnige en onzinnige gesprekken.

Iedere dag op de universiteit werd afgewisseld met een dag in een gewone huisartsenpraktijk. Mijn collega's en assistentes uit de praktijk volgden mijn vorderingen met interesse, verbazing en ongetwijfeld ook af en toe geamuseerd. 'Wanneer is je proefschrift klaar?', werd na een klein jaar jaar een vaste terugkerende vraag, terwijl ik met mijn hoofd alleen nog bij de subsidie voor de hoofdstudie was. Het is nu af, en voor onze plannen en de samenwerking in de komende jaren is dat goed.

Van je familie moet je het hebben. In de laatste fase van de promotie kreeg ik van diverse kanten hulp. Mijn zus, Tiny Hobma, wist ondanks haar eigen drukte en beperkingen snel en goed commentaar te leveren op mijn engels, en gaf en passant nog algemene grammaticale tips en trucs. Mijn schoonvader, Sjef Meijers stak veel geduld en tijd in de omslag en vormgeving van dit boekje. Raf, bedankt voor jouw hulp bij de niet-wetenschappelijke maar minstens zo belangrijke afronding van dit project. Thuis, van Edith, Ernst en Lucas, kreeg ik altijd de ruimte die ik nodig had voor het project, of het nu was om te werken, of gewoon geestelijk afwezig te zijn ('gaat er eigenlijk wel eens een uur voorbij dat je niet aan je project denkt?').

En dan de echte inner circle als het om dit onderzoek gaat. Karin Aretz en Willy Dubois waren als onderzoeksassistente aan het project verbonden. Willy hield zich bezig met het verwerken van toetsen en het maken van mooie feedbackrapporten, en het ontwikkelen van speciale software daarvoor. Toen ging ze op vakantie, en kwam niet meer terug. Willy had toen al iets neergezet waarmee het project moeiteloos verder kon; haar dood door het verschrikkelijke ongeluk blijft moeilijk te bevatten. Karin zette de databestanden op en samen deden we de analyses. Karin, je ervaring met dit soort projecten en je kennis waren een voorwaarde om op een goede manier door deze fase heen te komen, en het was zeer prettig met je samen te werken. Bedankt.

Cees en Richard, pas toen ik in het buitenland op congres was kreeg ik in de gaten dat ik geen alledaagse promotoren had. Hoewel dat tijdens de rit niet altijd een voorrecht leek, kijk ik er nu, op het eind, wel zo op terug. Jullie benadering was zeer verschillend. Cees, sterk georiënteerd op 'leren', altijd in voor discussie en de kracht van het argument, Richard, gericht op de weerbarstige praktijk en op het hoogst haalbare resultaat. De inhoud van jullie leerstoelen en werkstijl hebben zeker iets met elkaar te maken. Tenslotte 2 mensen met wie ik in de afgelopen jaren wel zeer intensief heb samengewerkt. Paul, we hebben heel veel samen gedaan de afgelopen jaren. Dat we daarmee de komende tijd doorgaan zegt denk ik genoeg. Paddy, jij was de juiste vrouw op de juiste plaats, zonder jouw werkdrijf en perfectionisme was dit proefschrift er niet geweest. Hoewel jullie het weten, toch: ik heb het zeer gewaardeerd, bedankt.

CURRICULUM VITAE

Sjoerd Hobma werd geboren in Bolsward op 20 juni 1964. In 1983 behaalde hij zijn VWO-diploma aan de scholengemeenschap Canisius College-Mater Dei in Nijmegen, en begon hij de studie geneeskunde in dezelfde stad. In 1991 behaalde hij zijn artsexamen. Van 1991 tot 1995 werkte hij als arts-assistent interne geneeskunde in het Canisius-Wilhelmina Ziekenhuis en het Radboud Ziekenhuis in Nijmegen. Vervolgens was hij een jaar werkzaam bij de GGD Oostelijk Zuid Limburg als sociaal en forensisch geneeskundige. In 1997 startte hij de huisartsopleiding aan de Universiteit Maastricht. Vanaf 1999 is hij in deeltijd werkzaam als huisarts in Sittard, en in deeltijd verbonden aan de vakgroep huisartsgeneeskunde in Maastricht, aanvankelijk als onderzoeker, en vanaf 2004 als staflid van de huisartsopleiding. Hij is getrouwd met Edith Meijers, en heeft twee geweldige zonen: Ernst en Lucas.