Unlimited exposure. The patient mix of GP trainees and their trainers: gaps, disparities, and active steering

de Jong, J.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: http://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
The patient mix of GP trainees and their trainers: gaps, disparities, and active steering

Jip de Jong
UNLIMITED EXPOSURE

THE PATIENT MIX OF GP TRAINEES AND THEIR TRAINERS: GAPS, DISPARITIES, AND ACTIVE STEERING

JIP DE JONG
“Why should I struggle through hundreds of pages of fabrication to reach half a dozen very little truths?’
‘For fun?’
‘Fun!’ He pounced on the word.
‘Words are for truth. For facts. Not fiction.’
— John Fowles, The Magus
UNLIMITED EXPOSURE

THE PATIENT MIX OF GP TRAINEES AND THEIR TRAINERS: GAPS, DISPARITIES, AND ACTIVE STEERING

ACADEMISCH PROEFSCHRIFT

ter verkrijging
van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. D.C. van den Boom
ten overstaan van een door het college
voor promoties ingestelde commissie,
in het openbaar te verdedigen
in de Aula der Universiteit
op woensdag 31 oktober 2012,
te 11:00 uur

door
JIP DE JONG
geboren te Amsterdam
PROMOTIECOMMISSIE

Promotor
Prof. dr. M. Wieringa-de Waard

Co-promotor
Dr. M.R.M. Visser

Overige leden
Prof. dr. M.J. Heineman
Prof. dr. H.E. van der Horst
Prof. dr. A.D.C. Jaarsma
Dr. E.P. Moll van Charante
Dr. W. de Ruijter
Prof. dr. F.G. Schellevis

Faculteit der Geneeskunde
CONTENTS

7 Chapter 1 General Introduction

25 Chapter 2 A systematic review on the relationship between patient mix and learning in work-based clinical settings. Submitted for publication

53 Chapter 3 Which barriers affect morbidity registration performance of GP trainees and trainers? Submitted for publication

71 Chapter 4 Opening the black box: The patient mix of GP trainees British Journal of General Practice 2011;61:e650-e657(8)

91 Chapter 5 Exploring differences in patient mix in a cohort of GP trainees and their trainers. BMJ Open 2011;1:e000318

113 Chapter 6 Who determines the patient mix of GP trainees? The role of the medical receptionist. Family Practice 2011;28:287-93

131 Chapter 7 Steering the patient mix of a general practitioner trainee is not as easy as it seems. Results of a randomized controlled intervention. Provisionally accepted

155 Chapter 8 General Discussion

179 Chapter 9 Summary

187 Chapter 10 Samenvatting

197 Dankwoord

205 Over de auteur
CHAPTER 1

General Introduction
General Introduction

This thesis concerns a series of studies into the patient mix of general practitioner (GP) trainees. In this introduction, first background information on the patient population of GPs and the GP specialty training is presented. Subsequently, the term ‘patient mix’ is defined and an important learning theory, deliberate practice, denoting the significance of patient mix, is presented. The need for obtaining a clear picture on the patient mix of GP trainees is discussed as is the identification of areas in which the patient exposure of GP trainees is low. The possible causes for low exposure are addressed. The possibility of actively steering the patient mix is introduced. Finally the aim, methods, and outline of the thesis are presented.

THE PATIENT POPULATION OF GPs IN THE NETHERLANDS

The Dutch health-care system is unique in the world. GPs in the Netherlands have a gatekeeper role, meaning patients cannot consult a specialist without being referred by their GP. A basic health-care insurance that covers care delivered by a GP is mandatory for all inhabitants. Dutch GPs therefore provide first-line, primary care to practically all inhabitants and, consequently, are confronted with a broad range of complaints and diseases (‘patient mix’). Out-of-hours care is mostly organized in GP cooperatives. Dutch GPs see patients with all kinds and stages of acute and chronic medical conditions, psychological disorders, and social health problems. These factors contribute to the fact that primary care epidemiology is highly different from specialized (secondary) care. Patients are registered with a permanent GP and consequently the continuity of the relationship between patient and doctor is an important aspect. Most patients prefer to consult their ‘own’ GP or GP practice whenever a new health problem arises.¹

Due to the way in which the Dutch health-care system is organized, GPs see many different health problems and must therefore be competent in many areas. This is supported by the The Dutch College of General Practitioners (NHG) and the Dutch Association of General Practitioners (LHV) stating in ‘Health Care Supply General Practice 2009’ that GPs are capable of delivering a broad range of care.²

THE GP SPECIALTY TRAINING PROGRAMME

To become a fully licensed GP in the Netherlands, qualified medical doctors have to follow an intensive, 3-year training, at one of the eight GP specialty training institutes. In these 3 years, the competencies are acquired to work as a responsible,
independent health professional able to deal with the broad patient mix that is presented. Two of these 3 training years, the first and the third year, are spent in a GP training practice. Here GP trainees work under the supervision of a GP trainer. For this, GP trainees are matched with their trainers and consequently placed within the GP trainer’s practice. These GP trainers have at least 5 years’ experience as a GP and have worked for at least 1 year at their present location. To become a GP trainer, licensing is demanded, as is participation in a comprehensive training programme consisting of various didactic skills.

During the 3-year GP training course, GP trainees learn by receiving formal classroom education for 1 day a week at the training institute, and by self-study, but foremost by seeing patients in daily clinical practice under the supervision of their GP trainer. GP trainees hold daily surgery hours during which they see patients independently. During their surgery hours, the trainer is on stand-by should problems arise or should any form of assistance be required by the trainee. Coaching sessions are held on an almost daily basis.

**The problem**

Supervised work-based learning, gaining experience by seeing patients, is the core of the training. There is, however, no formal system to monitor patient mix, so the exact patient mix a GP trainee actually sees during surgery hours is not known. This means that the trainer and trainee as well as the training institute have no formal information about the number of patients, the gender and age distribution, or the types of patient health problems the trainees encounter. Both nationally and internationally (see Text box 1), the training institute and the GP trainer are regarded as being responsible for delivering a sufficient and adequate patient mix. For this, a clear view of the GP trainees’ patient mix is essential.

**Definition of patient mix and patient mix measures**

In order to obtain a view of the patient mix of GP trainees, a clear and workable definition of ‘patient mix’ has to be formulated. The first article with ‘patient mix’ in the title in Medline is a commentary by Edward Brandt Jr in 1974 on an article of McAllister and Dzur about the patient population in an acute medical care service. In this commentary, the author states that the ‘number of patients’ and the ‘types of medical problems’ are of prime importance and that clinical learning involves both quality and quantity. Of course, the issue of the health problems that
doctors in training are exposed to is much older, and numerous papers have been published following this quality and quantity or patient-volume and diagnosis-diversity approach.\textsuperscript{10-12}

In these papers, there is a large semantic overlap between the terms ‘clinical experience’, ‘clinical exposure’, ‘clinical encounters’ or ‘patient encounters’ and ‘case mix’ or ‘patient mix’. ‘Clinical exposure’ can be regarded as the umbrella term for clinical contacts of any kind. The term ‘patient mix’ inclines towards the description of the diversity of the exposure, focusing not only on variety or diversity, but also on quantity or volume. Because it is a more focused term than ‘clinical exposure’ or similar terms, it is the preferred term in this thesis.

Throughout this thesis patient mix is \textit{measured} and a simple definition of patient mix \textit{measures} (‘case mix’ in the publication) was formulated by Hutchinson.\textsuperscript{13} ‘A system of classifying ‘cases’ – patients, contacts, episodes, or visits – into groups which are similar according to some characteristic, such as diagnosis (e.g. International Classification of Diseases), treatment (e.g. OPCS operations codes), severity, potential for health-care improvements, or costliness.’

\textbf{DELIBERATE PRACTICE: THEORETICAL RELATIONSHIP BETWEEN PATIENT MIX AND LEARNING}

The way medical trainees acquire knowledge has been studied for many years and various theories have been developed. An important theory was postulated by K. Anders Ericsson, a Swedish psychologist and Professor of Psychology at Florida State University, USA.\textsuperscript{14} His ‘Deliberate Practice’ theory assumes that becoming an expert requires sufficient practical training in representative tasks and appropriate reflection which can be stimulated by feedback from coaches or trainers.\textsuperscript{15} He states that expert performance is different from normal performance and denies the more traditional theory assuming that achievable performance is limited by genetically determined talent. He argues that expert performance is reached by life-long deliberate training to improve performance in a specific domain and he describes training as a ‘highly structured activity explicitly directed at improving competence in a particular domain.’\textsuperscript{16} Deliberate practice is further specified by repetitive performance of intended cognitive or psychomotor skills in a focused domain, coupled with rigorous skills assessment, that provides learners with specific, informative feedback, resulting in increasingly better performance, in a controlled setting.\textsuperscript{15-17} Ericsson explicitly discussed the appropriateness of deliberate practice theory for the performance development of medical professionals.\textsuperscript{15}
Based on this, Duvivier et al.\textsuperscript{18} recently (2011) described medical training programmes as developed to overcome weaknesses and to improve competence. The level of competence must be monitored to provide cues for further improvement. Deliberate practice based (medical) training is not the repetition of activities but a focused approach aiming for well-defined learning goals.\textsuperscript{18} Within this framework, patient mix is an important training condition because it embodies the required representative tasks in the medical domain at issue onto which the desired competence can be practised. The patient mix offers different experiences on which reflection and assessment can be made by the trainee themself, by the trainer, or eventually, by an external preceptor.

\textbf{Text box 1.}
\textit{International recognition of the importance of an adequate patient mix}

Patient mix has gained emphasis in the diverse accreditation standards of several countries.

\textit{The World Federation for Medical Education (WFME), Postgraduate Medical Education WFME Global Standards for Quality Improvement.}\textsuperscript{5}

‘Training locations must have a sufficient number of patients and an appropriate case-mix to meet training objectives. The training must expose the trainee to a broad range of experience in the chosen field of medicine and, when relevant, include both inpatient and outpatient (ambulatory) care and on-duty activity. The number of patients and the case-mix should allow for clinical experience in all aspects of the chosen specialty, including training in promotion of health and prevention of disease.’

\textit{Liaison Committee on Medical Education (LCME).}\textsuperscript{6}

‘Each course or clerkship/clerkship rotation must identify any core experiences needed to achieve its objectives and ensure that students receive sufficient exposure to such experiences. Similarly, although the proportion of time spent in inpatient and ambulatory settings may vary according to local circumstances, in such cases the course or clerkship/clerkship rotation director must ensure that limitations in learning environments do not impede the accomplishment of objectives.’
The placements within a training programme, combined with a range of other learning opportunities must, together, provide GP Specialty Registrars with exposure to a range of patients, clinical problems, training environments and training opportunities sufficient to deliver the GP curriculum and so equip them for a career in independent practice.

The trainer ensures that the trainee is exposed to patient-based and non patient-based activities, resulting in the development of the desired competences.

- **Level II (beginner):** guides the volume of the patient supply for the trainees on the basis of their competences.
- **Level III (advanced):** takes action to adjust the patient supply where necessary in terms of quantity and quality. Encourages the trainees to adjust the patient supply themselves as much as possible, together with the medical receptionist.
- **Level IV (experienced):** takes a pro-active attitude in order to ensure that the trainees are provided with the patient supply that meets the requirements of the practical learning plan.

**THE EMPIRICAL RELATIONSHIP BETWEEN PATIENT MIX AND LEARNING**

In the 1960s, medical education emphasized clinical problem-solving abilities as the fundamental, underlying skill to solve clinical problems, sometimes at the expense of teaching clinical knowledge. In 1978, Elstein found that the performance of doctors was not consistent in different problems. This phenomenon was called ‘case-specificity’. According to Wimmers et al., case-specificity is not solely the result of content knowledge but also of level of experience and level of case difficulty. Current theories assume case-specific knowledge is of predominant importance, so a voluminous and diverse clinical experience is crucially important for competence development.
According to these theories and the implications of the deliberate practice framework, it can be assumed that GP trainees benefit most from the training when their patient mix requires them to exercise tasks that are representative of their later work as a GP. Their patient mix should therefore be comparable to those of trainers or to the patient mixes they are later exposed to as fully licensed GPs. Such a similar patient mix enables them to practice all their learning goals over the full required spectrum. Empirical evidence about the precise contribution of patient mix for learning in the medical field seems inconclusive. Moreover, differences in the supervisory quality were found to be more predictive of learning than differences in patient mix were.

GAPS AND DISPARITIES IN THE PATIENT MIX OF GP TRAINEES

Earlier research suggests that the patient mix of GP trainees differs from that of their trainers. Studies addressing this found that the trainees’ patient mix consisted of more minor illnesses and fewer psychosocial, chronic, and severe conditions compared with their trainers, demonstrating low exposure to important health conditions. It is questionable whether trainees can build sufficient competence in low exposure areas because there may be too little repetitive training and reflection opportunities. Other differences found in previous studies considered circulatory, metabolic, skin, eye, ear, and respiratory diseases. These studies, however, are relatively old, were small scale (n=8), or even case studies (n=1), and often covered short periods (≤4 weeks). A larger study over a longer period had not been done recently.

REASONS FOR THE DISPARITIES:

PATIENT ATTITUDE AND THE MEDICAL RECEPTIONIST

Various factors may contribute to the gaps and disparities in the patient mix of GP trainees. Some are associated with unchangeable factors, such as the geographic location of the training practice, its proximity to an emergency unit, and the gender and age distribution as well as the social class of the population. Patients’ preference for seeing their GP instead of the trainee can also cause gaps in the patient mix of the trainee. In a literature review by Bonney et al., four aspects that were relevant to the patients’ preference were distinguished. The first was patient acceptance of being treated by trainees. The acceptance of trainees is generally high. Patients expressed
an altruistic attitude to being involved in training the doctors of the future. Exceptions were older patients or patients with chronic disease or patients with personal or emotional concerns. The second aspect was the patient attitude towards continuity of care. Patients seeing trainees valued follow-up, involvement, and accessibility by their usual GP. The third aspect was patient trust. Trust levels for residents were reported to be high, but older patients had reduced trust in them.44 An association was found between continuity of care, patient trust, and satisfaction. Trainees spend a limited period in the training practice and therefore have limited time to build trust, and it is unlikely that they can establish the level of trust their trainers receive. Patients may have ‘institutional trust’44 which is trust in the medical facility carried over to include trust in the staff of the facility. The fourth aspect was the desire for meaningful communication.45 This concerns the problems with communication with the trainee, the understanding of the role of doctors in training, and the communication around transfer of care. Patients with chronic illnesses, in particular, were found to value the personal continuity with a regular GP.46;47

Whereas patient attitudes towards trainees has been extensively studied, the influence of the assigning behaviour of the medical receptionists48 on the patient mix has not. In the Dutch situation, patients are registered with a permanent GP and trainees mainly encounter patients registered with their trainer. Patients usually make appointments by telephone or at the desk and receptionists may ask for the reason for the consultation and any preference for a consultation with either the trainer or the trainee. With this information in mind, the receptionist assigns the patient either to the trainer or the trainee. Obviously, receptionists also consider the urgency of the problem in combination with the doctors’ schedules, existing arrangements regarding assigning, the availability of walk-in hours, and unbooked surgery for emergencies. Their considerations when assigning patients were unknown.

STEERING THE PATIENT MIX OF TRAINEES

In order to solve the gaps in GP trainees’ patient mix and eventually to reduce the disparities between trainees and trainers, several authors have suggested that more balance can be obtained by steering the patient mix in a desired direction.21-23;33;35;38 Steering patients can also be desirable when an individual trainee needs more exposure to meet a specific learning goal. Adam and Oswald proposed steering by instructing medical receptionists.33 This has several practical difficulties, for instance,
because patients are free to choose a doctor and, as stated before, do not always wish to consult trainees.49;50 Trainees and trainers have influence on the patient mix as well. For instance, they will ask patients to return for follow-up consultations, or they can start to hold surgery hours in an old people’s home or start consultation sessions for diabetic patients. The active steering of the patient mix of GP trainees by instructing medical receptionists, trainees, and trainers, whether this is possible, and whether this would have a beneficial effect on learning, had never been studied.

USING DATA DERIVED FROM ELECTRONIC PATIENT RECORDS TO OBTAIN A CLEAR VIEW OF THE PATIENT MIX OF GP TRAINEES AND THEIR TRAINERS
To obtain a detailed overview of the patient mix trainers and trainees are exposed to, data on patient contacts, such as the diagnosis, the contact type (consultation, home visit), and age and gender of the patient are needed. These data can be obtained in interviews or in questionnaires, or recorded in a hand-written, optically scanned, or electronic logbook. A logbook is usually a card or a paper or electronic form which is completed after a patient contact. In a recent review, Denton concludes that logbooks currently in use have insufficient reliability or validity.51 Other authors found electronic logbooks to be more effective and accurate52 than hand-written logbooks, because the latter are more prone to underreporting.53;54 Data extractions from electronic patient record (EPR) systems have been proven to be valid for providing insight into the patient mix of medical students and GP trainees.52;55-58 They also respond to almost all of Denton’s recommendations for the ideal logbook,51 apart from the fact that these recommendations were formulated for students in a clerkship.

ICPC CODING
In the Netherlands, most GPs keep detailed EPRs, including a standardized system of diagnosis codes: the International Classification of Primary Care (ICPC). Stimulated by the NHG, all EPR systems now include the ICPC-1 codes. This allows uniformity in descriptions of the patient mix. The World Health Organization (WHO) has accepted ICPC-2 within the Family of International Classifications (WHO FIC) and users may use it as a diagnostic classification for primary care or general practice, wherever applicable.59 The differences between ICPC-1 and ICPC-2 with regard to diagnosis codes are minor. ICPC-2 has never been implemented in the Netherlands.
The validity of morbidity coding has been extensively studied.\textsuperscript{60-65} Comparison of 52 British studies in a systemic review by Thiru et al. was hampered due to the lack of standardized methods for assessment of EPR data quality.\textsuperscript{60} However, the studies consistently reported completeness, indicating that the EPR data were valid. Another review by Jordan et al.\textsuperscript{66} gives an overview of the completeness and correctness of computerized general practice medical records in 24 studies; in this review, a consistent finding was that the quality of morbidity recording varied, probably caused by differences in the distinctiveness of diagnosis.

In the GP specialty training, trainees are taught that providing each contact with an ICPC code is the desired standard. It may be assumed that the completeness of the ICPC codes in the trainees’ EPR contacts are high. Still a small but meaningful part of the contacts may not have been coded, which may, even nowadays, in an educational setting, account for structural bias of the patient mix description. The coding performance of the trainees and the trainers is not known, neither are their opinions about coding, or the barriers they encounter.

THE AIM OF THE PRESENT STUDY
The aim was to answer the following research questions.

1. What is known about the relationship between patient mix and learning in work-based clinical settings?
2. What is the relationship between the barriers to ICPC coding of GP trainees and trainers and their self-reported and actual coding performance?
3. What patient mix do GP trainees encounter in GP training practices and what are the differences in patient mix between first-year and third-year GP trainees?
4. What are the differences in patient mix between trainees and trainers?
5. Which assigning behaviour do receptionists report?
6. Is it possible to intentionally steer trainees’ patient mix?
7. Does greater exposure contribute to better learning?

METHODS
The CASANOVA (CASe AssigNment fOr GP VocAtional training) project was started in a cohort of 73 training practices in the Netherlands. In these training practices, the patient mix was assessed by EPR data extractions. First a pilot was done in Almere, in 2007, in 16 practices which learned that data processing by using the statistical modules within the EPR did not give the desired result and data
processing software outside of the EPRs had to be developed. Data were collected between March 2008 and February 2010. Sixty-two training practices were included for a full year, and 11 were included for 9 months, as this was the length of the trainees’ secondment. Each trimester, data were obtained by the GP trainers or by the researchers. The data from the first 6 months of each training practice were used for the baseline registration (Research Questions 3 and 4). The baseline registration of third-year trainees that were enrolled for 9 months was reduced to 3 months. Over the next 6 months, a randomized controlled trial was carried out addressing the effects of steering the patient mix. The practices were randomized into two arms. In the intervention arm, the patient mix was actively steered by instructing the medical receptionist, the trainer, and trainee. The instruction was to assign more skin diseases in the first steering trimester and more psychosocial conditions in the second. First-year trainees who were enrolled in the study for 9 months only participated in the first intervention period. Learning outcome was measured by study-specific, self-assessment questionnaires and an online knowledge test. In addition, the supervision quality of the trainer was measured. ICPC coding behaviour of the GP trainers and trainees was assessed in a study-specific questionnaire as was assigning behaviour of the medical receptionists. During the study project, a Best Evidence in Medical Education (BEME) systemic literature review was done on the relationship between patient mix and learning in work-based clinical settings.
OUTLINE OF THIS THESIS

Chapter 2 addresses the relationship between patient mix and learning in work-based clinical training settings in a systematic review (Question 1).

Chapter 3 addresses the barriers the GP trainees and trainers face in their daily clinical work as perceived to ICPC coding. Barriers that affect coding performance and specific situations prone to non-coding are identified. The results of a questionnaire survey are compared with the outcome of coding percentages measured in the EPR extractions (Question 2).

In Chapters 4 and 5, a cohort study of 73 training practices is described. Data were collected using the EPR of the GP training practices. Chapter 4 describes the patient mix of GP trainees and differences between the first and the third training year are investigated (Question 3); and Chapter 5 identifies the differences between the patient mix of GP trainees and their trainers. Gaps and disparities between these patient mixes are examined (Question 4).

Chapter 6 concerns the role of the medical receptionists in the composition of the patient mix of GP trainees and their trainers. The assigning behaviour of the receptionists was evaluated by qualitative (focus group) and quantitative (questionnaire) methodology. The aim was to obtain a better understanding of the role of receptionists with respect to the composition of the patient mix of the GP trainee (Question 5).

Chapter 7 describes a randomized controlled trial in which the patient mix of the GP trainees was actively steered by instructing the medical receptionists, the trainer, and trainee. Outcome measures were both the influence of steering on the patient mix (Question 6) and on learning outcomes (Question 7). To measure these learning effects, self-assessment questionnaires (self-efficacy) and an online knowledge test (psychiatric knowledge) were developed.

Chapter 8 discusses the results of the study. Implications for the GP specialty training and for future research are indicated.

Finally in Chapter 9, a summary is presented.
REFERENCES


2. Health Care Supply General Practice 2009 [Aanbod Huisartsengeneeskundige zorg 2009]. Dutch College of General Practitioners (NHG) and Dutch Association of General Practitioners (LHV).


6. Liaison Committee on Medical Education (LCME). Functions and Structure of a Medical School. Standards for Accreditation of Medical Education Programs Leading to the M.D. Degree. http://www.lcme.org/standard.htm


19. Wimmers PF, Splinter TA, Hancock GR,


A systematic review of the relationship between patient mix and learning in work-based clinical settings

Jip de Jong, Mechteld Visser, Nynke van Dijk, Cees van der Vleuten, Margreet Wieringa-de Waard

Submitted for publication
Abstract

**Background** Clinical workplace-based learning has been the means to becoming a medical professional for many years. The importance of an adequate patient mix for an optimal learning process is based on educational theory, and recognized by national and international accreditation standards. The relationship between patient mix and learning in work-based curricula as yet remains unclear.

**Aim** To review research addressing the relationship between patient mix and learning in work-based clinical settings.

**Method** The search was conducted across Medline, Embase, Web of Science, ERIC and the Cochrane Library from the start date of the database to July 2011. Original quantitative studies on the relationship between patient mix and learning were included. Methodological quality was assessed and results were extracted by two reviewers using pre-specified forms.

**Results** A total of 10,420 studies were screened on title and abstract. Of these, 298 articles were included for full-text analysis, which resulted in the inclusion of 22 papers. The quality of the included studies, scored with the MERSQI, ranged from 8.0 to 14.5. A positive relationship was found between patient mix and self-reported outcomes evaluating the progress in competence as experienced by the trainee, such as self-confidence and comfort level. Patient mix was also found to correlate positively with self-reported outcomes evaluating the quality of the learning period, such as self-reported learning benefit, experienced effectiveness of the rotation, or the instructional quality. Variables, such as supervision and learning style, might mediate this relationship. A relationship between patient mix and formal assessment has never been demonstrated.

**Conclusion** Patient mix is positively related to self-reported learning outcome, most evidently the experienced quality of the learning programme.
Introduction

Clinical workplace-based learning has played the leading role in educating medical professionals for many years. The importance of an adequate case or patient mix at that workplace for an optimal learning process is intuitively felt by many professionals and is recognized by several national\textsuperscript{1–3} and international\textsuperscript{4} accreditation standards. The World Federation for Medical Education emphasized this in its Global Standards for Quality Improvement, for Postgraduate Medical Education.\textsuperscript{4} It states ‘Training locations must have a sufficient number of patients and an appropriate case-mix to meet training objectives. The training must expose the trainee to a broad range of experience in the chosen field of medicine and, when relevant, include both inpatient and outpatient (ambulatory) care and on-duty activity. The number of patients and the case-mix should allow for clinical experience in all aspects of the chosen specialty, including training in promotion of health and prevention of disease.’

The idea that much experience is needed to become competent doctors fits theoretical frameworks. According to Ericsson,\textsuperscript{5} medical expertise develops by ‘deliberate practice’. He argues that expert performance is different from everyday performance, as it continues to improve as a function of more experience, coupled with deliberate practice. Expert performance is reached by actively acquiring and refining a cognitive mechanism to support continued learning and improvement. Becoming a medical expert thus requires engagement in practice and appropriate reflection, which can be stimulated by feedback from coaches or trainers.\textsuperscript{5} In addition, other frameworks for medical expertise emphasize the importance of clinical experience for learning, such as theories of cognitive structures\textsuperscript{6} and dual processing.\textsuperscript{7–9} The essence of these theories is that first conscious, intentional learning (deliberate practice) must be established before routines are automated. These automated routines are the basis of adequate medical handling.\textsuperscript{10} Within these frameworks, the experience needed is provided by an adequate patient mix, so patient mix is an important training condition. Following deliberate practice theory, a well-supervised patient mix can be assumed to substantially improve medical competence.

This review was carried out in order to evaluate whether this theory could also be confirmed by empirical evidence. Our primary aim was to systematically review research addressing the relationship between patient mix and learning in workplace-based clinical settings. Our secondary aim was to address the influence of additional variables (e.g. supervision and learning style) on this relationship.
For this purpose, ‘patient mix’ was defined as the number of patients and the types of medical problems presented to learners.\textsuperscript{11} It thus is regarded to consist of a number of patients presenting a certain diversity of diseases (Figure 1). Numerous publications in the medical educational field have followed this ‘volume and diversity’ approach.\textsuperscript{e.g. 12-14}

Figure 1. Patient mix model

Two-dimensional approach of patient mix. Consider a training programme in a large group of patients with the same diagnoses (A). Because there is no diversity, this cannot be considered to be a ‘patient mix’, but merely the training of a single skill or restricted clinical problem. Situation C expresses a situation in which a large diversity of skills and/or symptoms and diagnoses are theoretically possible, but very few or even no patients are present. This can be regarded as a patient mix, but an extremely meagre one – simply because there are no or very few patients. The line labelled ‘D’ expresses (an arbitrary) ‘cutoff point’ where the diversity is rich enough to start calling the population ‘patient mix’. All points labelled B are considered to be a patient mix.
Methods

ELIGIBILITY CRITERIA (LIST 1)
As we aimed to assess the strength of the relationship between patient mix and learning, only quantitative studies were included which were conducted with medical students/trainees at any level of the formal medical training/career. Patient mix volume, i.e. the quantity of patients encountered and the diversity of skills and/or symptoms and diagnoses had to be described. No simple cut-off for the width of this diversity could be given (Figure 1), but studies on the exposure to one restricted clinical problem or skill were excluded as they only described the volume of that skill or problem, and no diversity. Learning outcome had to be explicitly assessed. The relationship between patient mix and learning had to be quantified by statistical analysis.

INFORMATION SOURCES AND SEARCH STRATEGY
The search was conducted across five sources relevant to education in a clinical context: Medline, Embase, Web of Science, ERIC and the Cochrane Library. The search ran from the start date of the database to July 2011 and was not limited by language, geography, or research methodology. The search strategy was composed by a clinical librarian. The search strategy had to be able to find a ‘golden standard set’ in PubMed. This set contained thirty-eight articles, previously rated as being relevant to the review subject by the authors. The strategy was then translated to the search systems of the other databases.

STUDY SELECTION
Two authors (MW and JJ, or MV and JJ) individually and independently screened the titles and abstracts of all articles using the inclusion and exclusion criteria. Citations that were selected by one author but not by the other author were discussed in order to achieve negotiated consensus on inclusion or exclusion. In case of doubt or persisting disagreement in this phase, the article was included. The full text of all the potentially relevant articles was retrieved. The full-text articles were screened, again independently, by two authors, using the same criteria and were again compared. In case of disagreement, a decision on inclusion or exclusion was once more reached by negotiated consensus. Most studies that were excluded did not have an adequate description of patient mix or did not statistically address the relationship between
patient mix and learning. At each screening phase, each citation was marked as ‘yes’, ‘maybe’ or ‘no’. Inter-observer agreement of the screening phase was measured by Cohen's Kappa (linearly weighted). Manual searches were conducted across the citations of the papers that were coded, resulting in 17 more citations. These were screened by two authors, but none of them were included.

**List 1. Inclusion and exclusion criteria**

<table>
<thead>
<tr>
<th>Inclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Empirical, quantitative, educational studies with actual patients (no simulations).</td>
</tr>
<tr>
<td>• Study population: studies conducted with medical students/trainees at any level of the formal medical training/career.</td>
</tr>
<tr>
<td>• Patient mix, clinical encounters, or clinical experience in workplace-based learning had to be described. Patient mix had to be described in some detail, thereby addressing the volume as well as the diversity. Studies on the exposure to a restricted clinical problem or skill were excluded. Medical subspecialties were not excluded beforehand, as long as the patient mix was diverse.</td>
</tr>
<tr>
<td>• Learning outcome measures had to be described by self-reported measures, assessment by trainers, preceptors, or others, or by objective structured clinical examinations (OSCEs), multiple choice, or other written exercises.</td>
</tr>
<tr>
<td>• The relationship between patient mix and learning had to be quantified by statistical analysis.</td>
</tr>
<tr>
<td>• Studies in all languages were included.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Qualitative studies.</td>
</tr>
<tr>
<td>• Dental and veterinary curricula, any paramedical curricula, nursing curricula, physician assistant curricula, nurse practitioner curricula, dietetic curricula.</td>
</tr>
<tr>
<td>• Theoretical medical curricula (not work-based).</td>
</tr>
<tr>
<td>• Complementary/alternative medicine.</td>
</tr>
</tbody>
</table>
DATA EXTRACTION
A detailed data extraction form was developed using the Best Evidence in Medical Education (BEME)\textsuperscript{15} standard coding sheet and published reviews\textsuperscript{16,17} as a basis. All selected papers were coded by the authors in pairs (MW and JJ, or MV and JJ). This form contained a general description of the study design and participants, including the training level and the specialty training area (Table 1). In addition, patient mix instruments (e.g. electronic logbooks and questionnaires), volume/diversity descriptions (e.g. top 10 skills or diagnoses lists), learning outcome measures, and the relationship found between patient mix and learning were recorded. We also documented the highest level of the Kirkpatrick hierarchy\textsuperscript{18} on which learning outcomes were assessed. If additional variables were studied in relation to learning outcome (e.g. learning style and supervision), these were also recorded.

QUALITY OF THE STUDIES
To obtain an overview of the quality of the included studies, and thereby the validity of the outcomes, we assessed them with the recently developed Medical Education Research Study Quality Instrument (MERSQI).\textsuperscript{19} This instrument was chosen because, to our knowledge, it is the only instrument fitted for observational studies considering medical education. Two authors (JJ and MV) independently scored the quality of the included papers. In case of disagreement on item scores, a decision was reached by negotiated consensus.

DATA ANALYSES
The various ways in which patient mix was operationalized were categorized in equivalent approaches of volume and diversity descriptions. Learning outcomes were divided into self-reported outcomes and outcomes using formal assessments. The relationship between patient mix and learning is described in sections based on different learning outcomes, as this allowed for homogeneous reporting of results.
Results

SEARCH RESULTS
The search resulted in 11,098 titles. After removal of duplicates, 10,420 studies were reviewed based on title and abstract. Of these studies, 298 were identified as potentially relevant and included for full-text analysis, which resulted in the inclusion of 22 papers (Figure 2). The studies identified had insufficient homogeneous or quantitative data to allow meta-analysis or other formal synthesis. During screening of titles and abstracts, the inter-observer agreement kappa (linearly weighted) was 0.34 (MV–JJ) respectively 0.32 (MW–JJ). Table 1 provides a summary of study descriptions and outcomes. This table forms the basis for the inferences from the studies in the following paragraphs.

Figure 2: Flow chart

<table>
<thead>
<tr>
<th>Source</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pubmed</td>
<td>4480</td>
</tr>
<tr>
<td>Cochrane</td>
<td>1667</td>
</tr>
<tr>
<td>EMBASE</td>
<td>1899</td>
</tr>
<tr>
<td>Web of Science</td>
<td>1384</td>
</tr>
<tr>
<td>Eric</td>
<td>1668</td>
</tr>
<tr>
<td>Sum</td>
<td>11,098</td>
</tr>
<tr>
<td>Duplicate removal</td>
<td>- 678</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>10,420</strong></td>
</tr>
</tbody>
</table>

Phase 1
10,420 citations screened on title and abstract

Phase 2
298 articles were retrieved full text +17 citations via references (not included)

Phase 3

Phase 4
22 articles reviewed
METHODOLOGICAL QUALITY OF STUDIES

MERSQI sum-scores ranged from 8.0 to 14.5 (median 11.75). Ten studies reported on the internal structure of their outcome instruments by Cronbach’s Alpha (Table 1) or principal component analysis. Response rates, if presented, varied from 43% to 100%. Data analyses were appropriate in all but one study, and all were beyond the descriptive level. Four studies reported outcome at satisfaction level, whereas 10 studies measured knowledge and/or skills. Less than half of the studies (n=8) measured outcomes at the behavioural level. None of the included studies explicitly measured patient or health-care outcome.

TYPES OF STUDIES

In six studies, the mutual dependence of factors related to learning was addressed in a path analysis or structural equation modelling (SEM) (Table 1). Eight studies compared the patient mix of training sites and their contribution to learning. In three of these studies, similar sites were compared, three others compared academic vs. non-academic sites and two compared inpatients and outpatients. Four studies evaluated the learning effects of an intervention: the introduction of a rotation, a skill-training programme, identification of 10 preselected complaints, and a new internship. Two studies compared groups of medical students at a different phase of their training. In four studies, the groups and sites were homogeneous and no interventions were studied.

OPERATIONALIZATION OF PATIENT MIX

Patient mix was measured with various instruments (see Table 1), including questionnaires (n=11), interviews (n=1), and logbooks (n=13), the latter hand-written (n=9), electronic (n=2), or unspecified (n=2). Patient mix was mostly described as the variety of encountered skills and/or diagnoses. Skills were usually technical procedures, such as intubation or suturing. In some studies (n=6), the patient volume was the most pronounced patient mix characteristic. The diversity of the patient mix in these studies was often additionally addressed by one or two variables, but the reports lacked a detailed insight into the diversity of diagnoses. In most other studies, the distribution of encountered diagnoses and medical skills was presented. Several authors presented a top 10 or 20 of the conditions the students meet most frequently. This method was also used to compare the patient mix of different sites.
OPERATIONALIZATION OF LEARNING

Learning outcome measures can be divided into self-reported outcomes and formal assessments. Self-reported outcomes are used in ten studies, in five of these, the self-estimated competence was measured as self-confidence or comfort level. In five other studies, the quality of the learning experience or the educational profit of the experience at issue is asked for; such as the effectiveness, the learning benefit, or instructional quality of the rotation. Formal assessments were more diverse. Usually, knowledge or skills were tested; sometimes including clinical performance. Methods used included multiple choice examinations (MCQ) and other written examinations, clinical assessments, oral examinations, and OSCEs.

RELATIONSHIP BETWEEN PATIENT MIX AND LEARNING

The relationship was usually tested by comparing the volumes of diagnoses or skill-diversity, or by determining the loading of one patient mix variable to an outcome variable in a path analysis.

Outcomes based on self-reporting

In four studies, a positive relationship was found between patient mix and self-reported outcomes evaluating the progress in competence as experienced by the trainee, such as self-confidence and comfort level. By contrast, one study found no difference in confidence between residents in a traditional inpatient rotation and a new one in which experience in ambulatory settings was introduced. Patient mix was also found to correlate with self-reported outcomes evaluating the quality of the learning period, such as self-reported learning benefit, experienced effectiveness of the rotation, or the instructional quality.

Outcomes based on formal assessment

MCQ OR OTHER WRITTEN EXAMINATIONS

In one trial, students in an intervention group who encountered significantly more often 10 chief prerequisite complaints than the control group (31.8% vs 6%) outperformed the control group on a general knowledge examination (p=0.014). In six other studies, however, no relationship between patient mix and scores on MCQs or other written examinations was found.
OSCE
No association was found between patient mix and performance in four of the five studies using OSCE assessment.²⁰;³⁴;³⁶;³⁸ Fung et al. suggested that the time allotted for students to complete clerkships may not be sufficient to expose them to the number of patients needed to generate a significant effect on clinical performance.³⁷ In one study, the OSCE scores even seemed lower in students who attended a higher number of outpatient clinics than those attending fewer outpatient clinics, although experience with emergency admissions and obtaining feedback on these seemed to improve OSCE performance.³⁶ The authors concluded that the clinical skills were enhanced by an increased volume of some, but not all, clinical experience. Jolly et al. found that students scored higher on OSCEs if they examined patients on their own, if the objectives (presumably the objectives of the rotation, not reported) had been made clear, and a higher number of clinics were attended.³⁸

ORAL EXAMINATION, CLINICAL ASSESSMENT
Wimmers et al. found that an increased number of patient encounters did not directly lead to improved clinical performance as assessed by supervisors in 227 medical students,⁴¹ as was the case in two other studies.³⁶;³⁹ They did, however, find a strong relationship between number of patients and number of diseases encountered (r=0.89).⁴¹ Ahmed and Hughes, in contrast, found that exposure rates did correlate with the assessment grades awarded by clinical supervisors, but not with a written exercise (quiz) score.³⁵ Also Greenberg and Getson found a weak positive correlation between number of patients seen and the students’ clinical performance.⁵⁷

VARIABLES POTENTIALLY RELEVANT TO THE RELATIONSHIP BETWEEN PATIENT MIX AND LEARNING
Martin et al. found that trainees with a deep, strategic, and well-organized learning style reported significantly higher clinical exposure (combined score for three areas of clinical activity). The well-organized style was also associated with OSCE performance.²⁰ McManus et al. additionally found that the amount of knowledge gained from clinical experience was related to strategic and deep learning styles³⁹ as was success in a final examination: positive and significant correlations were found for deep and strategic learning, whereas surface learning correlated negatively.
In the path analysis presented by Wimmers et al.,\textsuperscript{41} supervision quality loaded on patient mix volume and on clinical competence. Hoifoidt et al., however, did not find supervision to load on patient mix volume or on (subjective) learning benefit\textsuperscript{40} while in the model of Van der Zwet et al., supervision loaded on both patient mix and instructional quality.\textsuperscript{24} In the study by Dolmans et al., a relationship was found between supervision and the effectiveness of a rotation.\textsuperscript{21} Also, a significant two-way interaction was found between patient mix and supervision; the latter more strongly influenced the effectiveness of the rotation than patient mix did. In another study, OSCE score also seemed to be ‘modified’ by the quality of the feedback.\textsuperscript{36}

Ahmed and Hughes found indications of a relationship between professionalism and both patient mix and composite assessment grades at the end of an attachment.\textsuperscript{35} Hoifoidt et al. described that amount of experience affected the quality of the learning environment which itself was related to the learning benefit.\textsuperscript{40}

In the study by Yu et al., the overall quality of the surgical clerkship, as perceived by students, was related to the number of cases seen, although no difference in learning outcome was found.\textsuperscript{34} Jolly et al. found that six of 43 questionnaire variables correlated with OSCE score. Two of these six can be considered to be related to the learning climate, namely ‘whether students examined patients on their own’ and ‘whether objectives were made clear’.\textsuperscript{38}

Discussion

In most studies dealing with the relationship between patient mix and student self-assessment (self-confidence, comfort level), indications of a relationship were found. The indications of positive relationships were stronger regarding the quality of the learning experience (learning benefit, instructional quality, or effectiveness of a rotation). Supervision quality seems to be a mediating factor, which was repetitively found to improve patient- or education-related outcome.\textsuperscript{42} This can be regarded to be consistent with the theory of deliberate practice.

The relationship between patient mix and learning outcome was not corroborated with formal assessment outcomes. All but one study dealing with MCQ or other written examinations failed to find any relationship between patient mix and MCQ or written examinations. All the studies relating patient mix to OSCE score
found no association, or under some conditions even a negative association. In one study, a correlation was found between exposure rate and clinical assessment grades, whereas three other studies did not find such a relationship.

The patient mix (also called ‘clinical exposure’ or ‘case mix’) in the articles we reviewed was mostly presented without definition. We found studies describing skills, diagnoses, treatments, or general ideas about patient mix, within different specialties and measured by logbook or questionnaires and presented differently, making the patient mix descriptions extremely heterogeneous. The heterogeneity we found is particularly interesting. In the light of the emphasis, adequate patient mix has gained in the diverse accreditation standards of several countries and internationally. Due to the heterogeneity, we had difficulty in finding a proper cut-off point for the number of diagnoses or skills that need to be engaged to fulfil the diversity inclusion criterion. This heterogeneity indicates the need for a discussion on the value of the concept. For instance, Berlowitz et al. stated that patient mix should describe how patients are distributed along characteristics that may affect specific outcomes of interest; he thereby stresses that the concept of patient mix in itself is not relevant. The fact that the patient mixes described in the reviewed studies are so diverse, may be partly because they are related to different outcomes in the different settings at different stages of education. It often seemed that the presented patient mix depended on the instrument the authors had to their disposal and not on study-specific operationalization of the desired patient mix of the attachment. Besides the relationship with the outcome, more clarity about the relationship between the diversity and volume aspects of patient mix might be strived for. In this review, we found operationalizations of patient mix that were fairly different in that respect, allowing for very few inferences between studies.

The learning outcome measures were classified into self-reported assessment and formal assessment. The precise description of the used formal assessment methods in the studies was often meagre; example questions or exercises were not found. The reliability of clinical assessments is questionable; subjectivity can be a problem. Pulito et al. found that direct observations of trainees interacting with patients occur too infrequently. Students prepare for assessments and their results may reflect their preparation more than their real competence. Terms such as OSCE or MCQ might suggest that similar instruments were used in different studies. However, the precise content and the number of stations or questions were found to differ, if mentioned at all.
We formulated six possible explanations that could explain why we found so little evidence for the relationship between patient mix and the results of formal assessment.

1. **Patient mix does not contribute to medical competence development.**
   This idea is highly unlikely, although theoretically possible. The positive relationship between patient mix and self-assessment outcome (compared with formal assessment) is not per se an indication of a relationship between patient mix and learning. The largely absent relationship between patient mix and formal assessment might indicate that ‘clinical experience without training increases confidence but not competence’[39;46;47] The idea that one becomes automatically more competent with increasing experience can be illusory.

2. **The relationship between patient mix and learning is more complex and many other variables play a role (such as supervision quality, learning style, learning environment, or professionalism).**
   Based on educational theory,[5;48] the importance of other variables in the relationship could be expected. In several studies,[21;24;41] supervision was found to be strongly related to learning outcome. It may, therefore, be seen as an important mediator. Supervision quality was, however, not described or measured in the majority of the included studies, and the potentially mediating effect may have been overlooked. This may have been the case with other variables related to learning as well.

3. **The time span covered in most studies was too short.**
   Current educational theories assume a general problem-solving ability, but case-specific knowledge is considered of predominant importance.[49] This means that competences do not transfer easily,[50;51] which implies that long exposure in many domains is essential for doctors to become fully competent. The time-span covered by most of the included studies, may have been too short to find positive results.

4. **The patient mix is inadequately measured.**
   Patient mix is usually described by encountered skills or diagnoses and in terms of volume and diversity. Other potentially relevant descriptors are the complexity in relation to the stage of learning and the learning value or benefit of cases. These aspects, with exceptions, are not usually described, so the validity of the instruments might have been imperfect.
Furthermore, in several studies, the patient mix was aggregated per training site and comparison was made between sites, not between students. Maybe this is a too coarse comparison to establish the relationship between patient mix and learning which can also be regarded as a limitation of our inclusion and exclusion criteria.

5. The validity of formal assessment is insufficient.  
In a systematic review, Hamdy et al. found only mild to moderate correlations between measurements obtained in medical schools and future performance in medical practice. The clinical validity of OSCEs was also questioned by McManus. OSCEs, MCQs and other assessments may not be appropriate for determining the specific contribution of patient mix on learning.

Self-reported outcome instruments are usually designed especially for the study, whereas the formal assessments used are commonly part of the standard assessment procedures. These are not tailored to the study question and may suffer from bias due to the preparation of students for assessments. The self-report instruments might therefore be a more appropriate fit for the research questions.

6. The quality of the studies was insufficient.  
The majority of the included studies had a single-group, post-test-only design which may be considered inferior to pre–post test designs (n=4) or trials (n=2). Several studies were merely a comparison of training sites or an evaluation of a new curriculum.32;35

LIMITATIONS  
Composing an efficient and sufficient search strategy is complex. Despite our attempts to sharpen the patient mix definition to an accurate and workable one, we made pragmatic choices. We were not able to formulate the exact border between ‘some disease (or skills)’ and ‘patient mix’. This resulted in low inter-rater agreement. A substantial number of papers were included or excluded based on negotiated consensus. This happened more in the beginning of the review process (Phase 2) than later (Phase 3). A minority of the studies included were intended to specifically explore the relationship between patient mix and learning for general educational theory purposes. Most of the studies concerned merely an evaluation of a programme or curriculum change, or were a comparison between training sites.

Many studies had to be excluded because they lacked a statistical analysis of the relationship between patient mix and learning outcome.
FUTURE DIRECTIONS
This systematic review emphasizes the problem with the description of ‘patient mix’. Despite its attention in international accreditation standards, the concept itself seems poorly defined. Educational research would benefit from a standardized approach in patient mix descriptions; volume can always be measured, but diversity should be explicated in relation to the outcome. Future studies should aim at addressing which parts of patient mix (e.g. volume and diversity) contribute to learning and which parts do not. A theoretical framework accounting for other relevant parameters in the relationship between patient mix and learning, such as supervision and learning style, may be helpful, and instead of using the standard assessment procedures, research question tailored, objective outcome should be developed.

Nearly all studies at hand ‘accepted’ the patient mix that was presented to the participants as a given factor. Interventions on the patient mix were indirect (curriculum change) active influencing of the patient mix was not found. It would be interesting to see what the effect of tailoring the patient mix to the specific learning goals and needs of individual students would yield.

To avoid bias due to preparation for an examination, study outcome (assessments) should be unobtrusive; for instance, assessments based on a random selection of routinely made video recordings could be considered. If, at second best, a traditional approach is chosen – similar to the designs we found, triangulation should be strived for. This can be reached by measuring study-tailored self-assessment, including the quality of the learning experience and self-confidence or alike, combined with formal assessment derived from knowledge assessment, and assessment of clinical competence. These studies are preferably done in multi-institution trials. An inquiry into the detailed aspects of patient mix, and the contribution of these aspects to learning is desirable. This may be done in a qualitative design; trainees and clinical teachers may be interviewed about their ideas of minimal or optimal patient volumes and spread of diagnosis diversity and their benefit for learning.

CONCLUSIONS
Patient mix, defined in terms of volume and diversity, is related to self-reported learning outcome, most evidently the experienced quality of the learning programme. A relationship between patient mix and the results of formal assessment has rarely been demonstrated. Not only supervision in particular, but also learning style seem mediating variables of the relationship between patient mix and learning.
<table>
<thead>
<tr>
<th>Author, Year of publication</th>
<th>Country</th>
<th>Design</th>
<th>Time span</th>
<th>Speciality</th>
<th>Nr. of participants analysed and Educational level participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ahmed, 1999 UK</td>
<td></td>
<td>Single group, post test</td>
<td>7 weeks</td>
<td>Paediatrics</td>
<td>226 Students</td>
</tr>
<tr>
<td>Boots, 2008 Australia</td>
<td></td>
<td>Pre-post test design with two groups</td>
<td>10 weeks</td>
<td>Internal medicine</td>
<td>220 Students and 174 Interns</td>
</tr>
<tr>
<td>Chatenay, 1996 Canada</td>
<td></td>
<td>Single cohort randomized to attend one out of four training sites. Post test only</td>
<td>10 weeks</td>
<td>Surgery</td>
<td>109 Students</td>
</tr>
<tr>
<td>Dolmans, 2002 NL</td>
<td></td>
<td>Observational cohort study</td>
<td>3-12 weeks per rotation</td>
<td>8 disciplines</td>
<td>1208 Residents</td>
</tr>
<tr>
<td>Duke, 2011 Canada</td>
<td></td>
<td>Non equivalent control group, pre-post test</td>
<td>4 weeks</td>
<td>Family medicine</td>
<td>79 Residents</td>
</tr>
<tr>
<td>Fung, 2007 USA</td>
<td></td>
<td>Retrospective single group, post test</td>
<td>Not specified (three clerk-ships)</td>
<td>Inpatient internal, ambulatory and family medicine</td>
<td>166 Third year students</td>
</tr>
<tr>
<td>Greenberg, 1999 USA</td>
<td></td>
<td>Single group, post test</td>
<td>8 weeks</td>
<td>Paediatrics</td>
<td>118 Students</td>
</tr>
<tr>
<td>Gruppen, 1993 USA</td>
<td></td>
<td>Single group, pre-post test</td>
<td>1 month</td>
<td>Internal medicine</td>
<td>43 Third year students</td>
</tr>
<tr>
<td>Hoifodt, 2004 Norway</td>
<td></td>
<td>Cross-sectional</td>
<td>4 months</td>
<td>Psychiatry</td>
<td>85 Preregistration house officers</td>
</tr>
<tr>
<td>Jacobson, 1998 USA</td>
<td></td>
<td>Observational cohort study</td>
<td>12 weeks</td>
<td>Internal medicine</td>
<td>43 Students</td>
</tr>
<tr>
<td>Jolly, 1996 UK</td>
<td></td>
<td>Single group, post test</td>
<td>4 + 16 weeks</td>
<td>Pathology, then medicine and surgery</td>
<td>152 Clinical students</td>
</tr>
<tr>
<td>Lampe, 2008 USA</td>
<td></td>
<td>Non-randomized trial</td>
<td>6 months</td>
<td>Emergency medicine</td>
<td>37 Senior medical students</td>
</tr>
<tr>
<td>PM instrument</td>
<td>PM description (Volume/diversity)</td>
<td>Learning Instrument (self-reported measure is specified, Cronbach’s alpha if reported)</td>
<td>Relation patient mix and learning</td>
<td>MERQ2I</td>
<td>Highest Kirkpatrick level</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>----------------------------------</td>
<td>--------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>Change in exposure from year 1 to year 2 in 42 conditions and 20 skills</td>
<td>MCQ† and other written exam (0.67–0.71) General and clinical competence assessment</td>
<td>Clinical experience did not relate to MCQ/written exam score</td>
<td>I3</td>
<td>3</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>Encounters (ordinal scale) of 15 specified skills</td>
<td>Questionnaire, self-confidence</td>
<td>There may be a relation between skill exercise and confidence</td>
<td>9.5</td>
<td>2</td>
</tr>
<tr>
<td>Hand written logbook</td>
<td>Put volume of elective/ER admissions, operations, scrub, outpatient clinics and procedures.</td>
<td>MCQ† (0.67), OSCE* (0.48), Clinical performance</td>
<td>Skills were enhanced by increased volume of some but not all clinical experience. Complex relation between feedback and OSCE* performance. Quality of feedback seems to mediate this relationship</td>
<td>I4</td>
<td>3</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>One variable composed of 3 questionnaire items (sufficient—patients, diagnostic variety and—patients independently dealt with)</td>
<td>Self-reported effectiveness questionnaire</td>
<td>Self-perceived effectiveness depends on Patient mix and supervision. Supervision more strongly influences effectiveness when patient mix is limited</td>
<td>I0</td>
<td>1</td>
</tr>
<tr>
<td>(Unspecified) Logbook</td>
<td>- Percentage of students that attended clinics of 18 specialties - Nr of outpatient clinics attended</td>
<td>MCQ† knowledge test Self-confidence questionnaire Oral and written feedback assessment</td>
<td>No difference between ambulatory and inpatient sites</td>
<td>I2.5</td>
<td>2</td>
</tr>
<tr>
<td>Electronic (PDA**) logbook</td>
<td>Number of patients in 6 diagnostic categories</td>
<td>OSCE* (0.34–0.65)</td>
<td>No relation between patient exposure and OSCE* score</td>
<td>I2</td>
<td>3</td>
</tr>
<tr>
<td>Hand written logbook</td>
<td>Number of patients in 4 diagnostic domains</td>
<td>Clinical performance, case presentation and NBME†</td>
<td>No relation between volume and exam score</td>
<td>II.5</td>
<td>2</td>
</tr>
<tr>
<td>Hand written logbook</td>
<td>Number of patients, Number of patients that encountered 20 diagnostic categories Top 19</td>
<td>Written exam</td>
<td>No correlation between the students levels of experience and knowledge</td>
<td>I2</td>
<td>2</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>Nr of subjects having experience in 12 psychiatric skills</td>
<td>Questionnaire, Subjective learning benefit (0.84)</td>
<td>Subjective learning benefit was related to amount of experience, competence and formal teaching programme. Supervision and previous experience had no impact on subjective learning</td>
<td>II.5</td>
<td>3</td>
</tr>
<tr>
<td>Hand written logbook</td>
<td>Top 10 Top 10 -Percentage of encounters of 6 diagnostic categories and a skill</td>
<td>Categorization of self-reported learning points</td>
<td>In and outpatient encounters differed. Learning differences between in- and outpatient apply to pathophysiology, evaluation/work-up and patient education/counselling.</td>
<td>8.5</td>
<td>2</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>Questionnaire responses handling patient volume and skills</td>
<td>OSCE* (0.69)</td>
<td>No relation between clinical experience and educational outcome</td>
<td>I3.5</td>
<td>3</td>
</tr>
<tr>
<td>Electronic (PDA**) logbook</td>
<td>Nr of students that met the pre-specified target complaints.</td>
<td>MCQ† and written exam</td>
<td>Group seeing a required number of representative patients showed better knowledge</td>
<td>I2.5</td>
<td>2</td>
</tr>
<tr>
<td>Author, Year of publication</td>
<td>Country</td>
<td>Design</td>
<td>Time span</td>
<td>Speciality</td>
<td>Nr. of participants analysed and Educational level participants</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------</td>
<td>--------</td>
<td>-----------</td>
<td>------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td>Martin, 2000</td>
<td>UK</td>
<td>Single group, post test</td>
<td>1 year</td>
<td>Medicine and Surgery</td>
<td>194 (150 returned learning style form) Students</td>
</tr>
<tr>
<td>Mcleod, I997</td>
<td>Canada</td>
<td>Observational cohort study</td>
<td>8 week</td>
<td>Internal medicine</td>
<td>40 Residents and 29 clinical clerks</td>
</tr>
<tr>
<td>McManus, I998</td>
<td>UK</td>
<td>Prospective study of two cohorts assessed at application to med school and at the end of their final year</td>
<td>About 5 years</td>
<td>Undergraduate curriculum, not otherwise specified</td>
<td>684 Students (1st cohort 301/ 2nd 383)</td>
</tr>
<tr>
<td>Nomura, 2008</td>
<td>Japan</td>
<td>Pre-post test design with non-equivalent control group</td>
<td>2 years</td>
<td>Multidisciplinary</td>
<td>2474 before +166 after Postgraduate 'residents' without clinical experience</td>
</tr>
<tr>
<td>O'Hara, 2002</td>
<td>USA</td>
<td>Single group, post test</td>
<td>4 weeks + 4 days</td>
<td>Women health care in a family medicine clerkship</td>
<td>445 Students</td>
</tr>
<tr>
<td>Saywell, 2002</td>
<td>USA</td>
<td>Single group, post test</td>
<td>4 weeks</td>
<td>Musculoskeletal medicine in a family medicine clerkship</td>
<td>445 Third year students</td>
</tr>
<tr>
<td>Schwiebert, I993</td>
<td>USA</td>
<td>Single group, post test</td>
<td>1 month</td>
<td>Family medicine</td>
<td>185 Third year students</td>
</tr>
<tr>
<td>Wimmers, 2006</td>
<td>NL</td>
<td>Single group, post test</td>
<td>12 weeks</td>
<td>Internal medicine</td>
<td>152 Students</td>
</tr>
<tr>
<td>Yu, 2011</td>
<td>New Zealand</td>
<td>Single group, post test</td>
<td>6 weeks</td>
<td>Surgery</td>
<td>166 Fourth year students</td>
</tr>
<tr>
<td>vd Zwet, 2010</td>
<td>NL</td>
<td>Single group, post test</td>
<td>10 weeks</td>
<td>General Practice</td>
<td>284 Fifth year students</td>
</tr>
</tbody>
</table>

* Objective Structured Clinical Examination
† Multiple Choice Questions
‡ National Board of Medical Examiners
** Personal Digital Assistant
<table>
<thead>
<tr>
<th>PM instrument and description</th>
<th>Learning Instrument (self-reported measure is specified, Cronbach’s alpha if reported)</th>
<th>Relation patient mix and learning</th>
<th>MERQSI</th>
<th>Highest Kirkpatrick level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Questionnaire</td>
<td>OSCE* (0.70)</td>
<td>No association between clinical experience and OSCE* score. Positive association clinical experience with learning style.</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>Interview, (unspecified) logbook, Questionnaire</td>
<td>Relevance for learning Questionnaire / interview</td>
<td>Inpatient based experience is better than ambulatory care experience for learning</td>
<td>8.5</td>
<td>2</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>MCQ and other written exam</td>
<td>No association between clinical experience and exam score. Study habits predict examination performance</td>
<td>14.5</td>
<td>3</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>Questionnaire self confidence</td>
<td>Clinical experience and confidence levels improved, especially at university hospitals</td>
<td>10.5</td>
<td>2</td>
</tr>
<tr>
<td>Hand written logbook</td>
<td>Questionnaire comfort level</td>
<td>Relationship between experience and comfort level between some diagnostic categories</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>Hand written logbook</td>
<td>Questionnaire comfort level</td>
<td>Relationship between experience and comfort level between some diagnostic categories</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>Hand written logbook</td>
<td>Written exam, oral exam</td>
<td>Slight differences between university and private practice in patient mix, but no difference in results on oral and written exam.</td>
<td>11.5</td>
<td>2</td>
</tr>
<tr>
<td>Hand written logbook</td>
<td>Combined clinical performance assessment and oral examination (0.67)</td>
<td>An increased nr. of patient encounters did not (directly) lead to improved competence. Quality of supervision indirectly had impact on student learning and the nr. of patient encounters</td>
<td>13</td>
<td>3</td>
</tr>
<tr>
<td>Questionnaire and hand written logbook</td>
<td>Clinical assessment, Critical Appraised Topic and OSCE* (0.69-0.74)</td>
<td>Heterogeneity of clinical exp from sites did not translate into heterogeneity of learning outcomes</td>
<td>13.5</td>
<td>2</td>
</tr>
<tr>
<td>Questionnaire</td>
<td>Questionnaire, instructional quality</td>
<td>Supervision and patient mix load on instructional quality</td>
<td>10</td>
<td>1</td>
</tr>
</tbody>
</table>
REFERENCES

1. Liaison Committee on Medical Education (LCME). Functions and Structure of a Medical School. Standards for Accreditation of Medical Education Programs Leading to the M.D. Degree. http://www.lcme.org/functions2010jun.pdf


15. BEME Collaboration. http://www2.warwick.ac.uk/fac/med/beme/


Appendix 1. Search strategies

1. PUBMED
   
   **Set 1: patient mix**
   

   **Set 2: learning**
   

   **Set 3: population**
   

2. EMBASE
   
   **Set 1: patient mix**
   
   case mix/ or case mix*.ti,ab. or casemix*.ti,ab. or diagnosis related group/ or diagnosis related group*.ti,ab. or clinical exposure*.ti,ab. or clinical encounter*.ti,ab. or clinical experience*.ti,ab. or patient mix*.ti,ab. or logbook*.ti,ab. or consultation.ti,ab. or selected conditions.ti,ab. or disease management.ti,ab. or clinical method*.ti,ab. or (diagnosis adj1 cluster*).ti,ab. or (distribution adj2 patient*).ti,ab

   **Set 2: learning**
   
   curriculum/ or curricul*[ti,ab. or exp Medical Education/ or medical education.ti,ab. or exp Clinical Competence/ or clerkship*[ti,ab. or trainee*[ti,ab. or training.ti,ab. or work based learning*[ti,ab. or (residency or resident*).ti,ab.

   **Set 3: population**
   
   exp Teaching Hospital/ or teaching hospital*[ti,ab. or exp medicine/ or exp Primary Health Care/ or student*[ti,ab. or practice.ti,ab.
3. COCHRANE LIBRARY

#1 (case mix*):ti,ab,kw
#2 (casemix*):ti,ab,kw
#3 (diagnosis related group*):ti,ab,kw
#4 MeSH descriptor Diagnosis-Related Groups explode all trees
#5 (clinical exposure*):ti,ab,kw
#6 (clinical encounter*):ti,ab,kw
#7 (clinical experience*):ti,ab,kw
#8 (patient mix*):ti,ab,kw
#9 (logbook*):ti,ab,kw
#10 (consultation):ti,ab,kw
#11 (selected conditions):ti,ab,kw
#12 (diseases management):ti,ab,kw
#13 (clinical method*):ti,ab,kw
#14 (diagnosis cluster*):ti,ab,kw
#15 (distribution patient*):ti,ab,kw
#16 (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15)
#17 MeSH descriptor Curriculum explode all trees
#18 (curricul*):ti,ab,kw
#19 MeSH descriptor Education, Medical

4. ERIC

1 exp "Case Method (Teaching Technique"/
2 exp Clinical Experience/ or clinical exposure.mp.
3 clinical encounter.mp.
4 logbook*.ti,ab.
5 exp Patients/
6 (case mix* or casemix*).ti,ab.
7 1 or 2 or 3 or 4 or 5 or 6
8 exp Curriculum/
9 curricul*.ti,ab.
10 exp Medical Education/
11 medical education.ti,ab.
12 clinical competence.mp.
13 exp "Clinical Teaching (Health Professions")/
5. WEB OF SCIENCE
Title=(“family practice” OR “general pract*” OR “family medicine” OR “primary care” OR “internal medicine” of psychiatr* OR “hospital*” or surgery) AND Title=(curriculum or training* OR trainee* OR clerks* OR residen* OR education* OR learn* OR medical student* OR internship* OR work based learning) AND Title=(“case mix” OR “casemix” OR “experience*” or disease* OR logbook* OR “patient mix” OR examination* OR patient* OR diagnos* OR condition*)
Which barriers affect morbidity registration performance of GP trainees and trainers?

Jip de Jong, Mechteld R.M. Visser, Margreet Wieringa-de Waard

Submitted for publication
Abstract

**Background** ICPC coding percentages in GP specialty training are high, but not 100%, indicating barriers against coding still exist, possibly influencing the validity of EPR-based data.

**Objective** To study the relationship between barriers to ICPC coding of GP trainees and trainers and their self-reported and actual coding performance.

**Methods** A questionnaire was developed, and returned by 71 (of 73, 97%) GP trainees and 103 (of 108, 95%) GP trainers, affiliated to the GP Specialty Training of the Academic Medical Center, University of Amsterdam. Their barriers to ICPC coding and self-reported coding performance were compared with EPR-derived data extractions that were collected during one year.

**Results** Mean coding percentages were 88.3 (SD=11.5) and 82.3% (SD=19.0) (trainees/trainers). Most participants reported always registering ICPC codes for consultations and home visits, specifically in those situations pre-specified in the questionnaire. Telephone consultations, repeat prescriptions and administrative actions were coded less frequently. Most participants never or rarely experienced coding barriers, an exception being ‘insufficient refinement of the ICPC system’. Most motivation and ICPC-related barriers correlated with self-reported and actual coding performance. ‘ICPC coding is unpleasant to use’ predicted the trainees’ coding percentage. This also predicted the trainers’ coding percentage, as did ‘no personal gain from ICPC’ and ‘coding is difficult’.

**Conclusion** Coding percentages can be improved by increasing motivation and by making ICPC coding more pleasant to use. As no other specific areas were prone to non-coding, EPR-derived data seem biased by non-coded telephone consultations only.
Introduction

Aggregated data based on medical classifications or clinical coding systems are important resources for clinical decision making and decision support, audit, governance, research, education, training\(^1\) establishing national or local morbidity figures\(^2\) and epidemic surveillance.\(^3\text{-}^5\) The benefits of the various functions in GPs electronic medical, health or patient records (EPRs) depend on good registration.\(^6\) Examples are: prescription modules\(^7\), contraindication surveillance\(^8\), an influenza vaccination module\(^9\), protocols for monitoring chronic patients and data used for annual reporting and monitoring patient mix of trainees.\(^10\text{-}^11\)

The validity of morbidity coding has been extensively studied.\(^12\text{-}^17\) A systemic review by Thiru et al. showed consistently high positive predictive values, indicating the validity of EPR data. However, comparing the 52 included studies was hampered due to the lack of standardized methods for assessing EPR data quality. Another review by Jordan et al.\(^18\) reviews the completeness and correctness of computerized general practice medical records in 24 studies in the UK. A consistent finding was that the quality of morbidity recording varied, probably through differences in diagnosis distinctions. A third review by Khan et al. analysing 49 papers showed that, apart from acute conditions, most diagnoses were well recorded.\(^17\)

The quality of coding is improved by training.\(^13\) However, GP’s did not always feel that morbidity registration was directly beneficial to their daily patient care. In a qualitative study using semi-structured interviews, De Lusignan explored barriers encountered by GPs regarding morbidity registrations.\(^19\) He found concerns that morbidity coding may jeopardise GPs’ relationships with their patients; sensitivity to diagnostic uncertainty and the potential impact of diagnostic labels. He also identified barriers to clinical coding in a narrative review,\(^20\) classifying them as technological; the limitations of the coding systems and terminologies and the skill gap in their use; integration of the clinical task; recording structured data in the consultation takes time and is distracting, individual; the level of motivation of primary care professionals and organizational; the priority within the organization. Other authors found similar results.\(^21\)

Several classification systems are used worldwide: ICD, ICPC, ICF, SNOMED and READ. Most Dutch GPs keep detailed EPRs, including a standardized system of diagnosis codes: the International Classification of Primary Care (ICPC).\(^22\) Stimulated by the Dutch College of General Practitioners (NHG), all EPR systems now
include the ICPC-1 codes. Developed by the ICPC Working Party and published in 1987 by the World Organisation of National Colleges, Academies (WONCA), the ICPC comprises 17 chapters (organ systems, general and social). The World Health Organisation includes the ICPC (Version 2) within its FIC (Family of International Classifications. The ICPC coding system allows the coding of both ‘general and unspecified’ diagnoses (e.g. headache, ankle complaints) and ‘specific’ diagnoses (e.g. pneumonia, appendicitis, hypothyroidism). The code differences between ICPC-1 and ICPC-2 are minor. ICPC is used in primary care in many countries, including Argentina, Australia, Denmark, Greece, Norway, The Netherlands, Spain, and others.

ICPC has proven valid for studying morbidity patterns, and useful for monitoring patient mix, in GP training. Patient monitoring based on EPR extractions provides a means of gaining insight in the clinical exposure, or patient mix, of GP trainees. Detailed EPR-based descriptions of the GP trainees’ patient mix can be used to determine if and where low exposure exists; for individuals or the entire group. Training programmes may be adapted accordingly. GP trainers can use detailed knowledge about the patient mix of their trainees to tailor this mix for specific learning objectives.

Coding is no longer seen as optional: trainees are taught that coding patient encounters is the desired standard. Since coding percentages in the GP training practices are high, many of De Lusignan’s former barriers seem to have been overcome. Yet some encounters are still not coded. If these encounters structurally concern specific contact forms or situations, data based on EPR extractions may be biased, thus invalidating these data for research purposes. This study investigates the coding performance of GP trainees and trainers, to provide more insight into any ICPC coding barriers they face in their daily work, and to identify specific situations prone to non-coding. The relationship between trainers/trainees’ barriers and their coding percentage is then addressed.
Methods

PARTICIPANTS
This study was part of a larger study exploring the patient mix of GP trainers and their trainees using EPR extractions. Questionnaires were sent to all 73 participating trainees (49 first-year and 24 third-year trainees) and to 108 trainers. All participants gave their informed consent.

STUDY SETTING AND PROCEDURES
This study was conducted in practices affiliated to the GP Specialty Training of the Academic Medical Center, University of Amsterdam. This training institute facilitates a 3-year training programme in which first and third year trainees are stationed in GP training practices of experienced GP trainers. In their second year trainees do clinical rotations. EPR extraction data was collected between March 2008 and January 2010. The questionnaire was sent at the end of this data collection.

QUESTIONNAIRE
The authors drafted a study-specific questionnaire, the face validity of which was discussed with 3 experts of the study advisory committee before it was amended and finalized. The questionnaire used five-point Likert scales and was based on literature and on our experience. The questionnaire comprised 30 questions in three sections. Firstly, participants were asked to estimate their own coding performance regarding face-to-face encounters, telephone consultations, home visits and administrative actions (processing mail, letters from specialists or laboratory/radiology reports). Then they could tick four pre-specified reasons why registering ICPC codes is difficult, or describe their reasons in free text. Secondly, they could indicate how far they agreed with statements describing potential barriers for ICPC coding (see Text box 1), covering motivation related, ICPC-system related or external factors. Thirdly, participants indicated how often they used ICPC codes for specified situations prone to non-coding, to provide information about situations in which ICPC coding is lacking. Participants could mention other non-coding situations in the free text space provided.
Text box 1. Example questions of the questionnaire

**Please indicate:**
- What proportion (approximately) of patient-contacts (face-to-face, telephone contacts and home visits together) you register with a code
- How often do you register regular face-to-face consultations with a code?
- How often do you register telephone consultations with a code?
  1 = hardly any consultations; 2 = less than half; 3 = half; 4 = more than half; 5 = almost all consultations

**Please indicate to what extent the next statements fit your working routine**
- Coding costs too much time
- Coding offers no personal gain
  1 = agree; 2 = partly agree; 3 = neutral; 4 = partly disagree; 5 = disagree

**Please indicate how often you register the situations mentioned in the next section with an ICPC code**
- Repeat prescriptions (not during a consultation)
- Medical examinations
- Complex problems
  0 = not applicable; 1 = never; 2 = seldom; 3 = sometimes; 4 = often; 5 = (nearly) always

**EPR EXTRACTED DATA**

To relate barriers and estimations of coding performance to actual coding behaviour, we calculated ICPC coding percentages using data-extraction software. During one year, raw data were extracted quarterly by either the GPs or the researchers and processed by multiple software routines. Decision rules were programmed to ensure logic and consistency in counting the ICPC codes. The resulting aggregated data were anonymized and e-mailed to the researchers. No information about individual patients could be recovered. The participants received feedback about their coding percentage after each data extraction.
DATA ANALYSES
Coding percentages were calculated by dividing the total number of ICPC diagnosis codes (face-to-face encounters, telephone consultations and home visits) in a trimester by that number plus the number of non-coded encounters, corrected by the individual mean number of ICPC diagnosis codes per contact. Repeat prescriptions and administrative actions were disregarded (SPSS18.0).

The overall mean coding percentage was computed only where data from at least two trimesters were present. Descriptive analyses were performed. Pearson correlations were calculated between the barriers and both self-reported performance (questionnaire-based) and actual coding performance (based on data extractions). Subsequently, in multivariate regression analyses (stepwise, probability of F-to-enter ≥0.05, to remove F ≥0.1) the same two outcome measures were predicted; separately for trainers and for trainees. The barriers were the predictors in each regression.

Results

RESPONSE RATE OF THE QUESTIONNAIRE AND COMPLETENESS OF DATA EXTRACtIONS
The questionnaire was returned by 71 trainees (97%) and 103 trainers (95%). EPR-data of only one trimester was extracted for one trainee and two trainers (migration to other practice, illness). Their data were not used in the correlation analyses or in the regression. Two trainers filled in the questionnaire but delivered no EPR-data to analyse.

CODING PERCENTAGE
Mean coding percentage of the trainees was 85.8% (SD=13.7, n=71) in the first trimester, 86.6% (SD=15.1, n=70) in the second trimester, 90.2 (SD=10.5, n=69) in the third and 90.9 (SD=10.5, n=58) in the fourth. The resulting overall mean percentage was 88.3 (SD=11.5, n=70). For trainers this was 78.8 (SD=21.8, n=99), 81.8 (SD=19.7, n=98), 85.2 (SD=17.8, n=96) and 86.6 (SD=16.1, n=82) respectively, resulting in an overall mean percentage of 82.3% (SD=19.0, n=99). Coding percentages rose significantly over time (p<0.05). Differences between trainers and trainees were non-significant.
BARRIERS AND GENERAL IDEAS ABOUT CODING

Both trainers and trainees reported registering ICPC codes for most patient encounters. Percentages concerning participants who indicated always registering ICPC codes for face-to-face encounters and home visits were particularly high (Table 1). Telephone consultations, and administrative actions were coded less frequently. Most participants did not (wholly) agree with the barrier statements, apart from the insufficient refinement of the ICPC system (Table 2).

DIFFICULTIES ENCOUNTERED WITH ICPC CODING

Pre-specified reasons why ICPC coding was found difficult were: unclear which code suits best 59%/57% (trainees/trainers), no suitable code 75%/52%, thesaurus works badly 11%/22%, the ‘.99 (category for specific diseases)’ codes are not refined enough 3%/17%. Reasons mentioned in free text (n=17) were time issues (n=9), problems with the ICPC coding system (n=3), practical problems (n=30) and ICPC provides no personal gain (n=1).

Table 1. Self-reported coding performance in percentages

<table>
<thead>
<tr>
<th>Self-reported coding</th>
<th>n</th>
<th>None</th>
<th>&lt; Half</th>
<th>Half</th>
<th>&gt; Half</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td>What proportion of all patient-contacts are you currently registering with a code?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trainee</td>
<td>71</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>18.6</td>
<td>81.4</td>
</tr>
<tr>
<td>Trainer</td>
<td>101</td>
<td>-</td>
<td>4.0</td>
<td>6.9</td>
<td>17.8</td>
<td>71.3</td>
</tr>
<tr>
<td>Face–to face consultations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trainee</td>
<td>69</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>10.1</td>
<td>89.9</td>
</tr>
<tr>
<td>Trainer</td>
<td>102</td>
<td>1.0</td>
<td>2.9</td>
<td>-</td>
<td>14.7</td>
<td>81.4</td>
</tr>
<tr>
<td>Telephone consultations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trainee</td>
<td>71</td>
<td>2.8</td>
<td>12.7</td>
<td>9.9</td>
<td>32.4</td>
<td>42.3</td>
</tr>
<tr>
<td>Trainer</td>
<td>102</td>
<td>8.8</td>
<td>12.7</td>
<td>8.8</td>
<td>21.6</td>
<td>48.0</td>
</tr>
<tr>
<td>Home visits</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trainee</td>
<td>71</td>
<td>-</td>
<td>-</td>
<td>4.3</td>
<td>10.0</td>
<td>85.7</td>
</tr>
<tr>
<td>Trainer</td>
<td>102</td>
<td>1.0</td>
<td>2.9</td>
<td>6.9</td>
<td>12.7</td>
<td>76.5</td>
</tr>
<tr>
<td>Administrative actions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trainee</td>
<td>71</td>
<td>20.0</td>
<td>21.5</td>
<td>21.5</td>
<td>20.0</td>
<td>16.9</td>
</tr>
<tr>
<td>Trainer</td>
<td>99</td>
<td>17.2</td>
<td>20.2</td>
<td>12.1</td>
<td>25.3</td>
<td>25.3</td>
</tr>
</tbody>
</table>
Table 2. Motivational, ICPC-related and external barriers in percentages

<table>
<thead>
<tr>
<th>Barriers</th>
<th>n</th>
<th>Agree</th>
<th>Partly Agree</th>
<th>Neutral</th>
<th>Partly disagree</th>
<th>Disagree</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Motivation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Takes too much time</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trainee</td>
<td>70</td>
<td>2.9</td>
<td>21.4</td>
<td>5.7</td>
<td>28.6</td>
<td>41.4</td>
</tr>
<tr>
<td>Trainer</td>
<td>102</td>
<td>3.9</td>
<td>22.5</td>
<td>12.7</td>
<td>24.5</td>
<td>36.3</td>
</tr>
<tr>
<td>No personal gain</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trainee</td>
<td>70</td>
<td>5.7</td>
<td>11.4</td>
<td>8.6</td>
<td>30.0</td>
<td>44.3</td>
</tr>
<tr>
<td>Trainer</td>
<td>101</td>
<td>5.9</td>
<td>13.9</td>
<td>14.9</td>
<td>24.8</td>
<td>40.6</td>
</tr>
<tr>
<td>No affinity with coding for research purposes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trainee</td>
<td>70</td>
<td>5.7</td>
<td>7.1</td>
<td>25.7</td>
<td>18.6</td>
<td>42.9</td>
</tr>
<tr>
<td>Trainer</td>
<td>100</td>
<td>8.0</td>
<td>10.0</td>
<td>20.0</td>
<td>14.0</td>
<td>48.0</td>
</tr>
<tr>
<td><strong>ICPC related</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICPC coding is unpleasant to use</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trainee</td>
<td>69</td>
<td>5.8</td>
<td>8.7</td>
<td>15.9</td>
<td>23.2</td>
<td>46.4</td>
</tr>
<tr>
<td>Trainer</td>
<td>101</td>
<td>4.0</td>
<td>11.9</td>
<td>18.8</td>
<td>20.8</td>
<td>44.6</td>
</tr>
<tr>
<td>ICPC Coding is difficult</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trainee</td>
<td>69</td>
<td>1.4</td>
<td>27.5</td>
<td>14.5</td>
<td>24.6</td>
<td>31.9</td>
</tr>
<tr>
<td>Trainer</td>
<td>102</td>
<td>2.9</td>
<td>14.7</td>
<td>22.5</td>
<td>28.4</td>
<td>31.4</td>
</tr>
<tr>
<td>ICPC is not sufficiently refined</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trainee</td>
<td>69</td>
<td>7.2</td>
<td>39.1</td>
<td>14.5</td>
<td>17.4</td>
<td>21.7</td>
</tr>
<tr>
<td>Trainer</td>
<td>102</td>
<td>4.9</td>
<td>28.4</td>
<td>24.5</td>
<td>21.6</td>
<td>20.5</td>
</tr>
<tr>
<td><strong>External</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Should be better paid</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trainee</td>
<td>69</td>
<td>7.2</td>
<td>4.3</td>
<td>15.9</td>
<td>8.7</td>
<td>63.8</td>
</tr>
<tr>
<td>Trainer</td>
<td>102</td>
<td>5.9</td>
<td>6.9</td>
<td>18.6</td>
<td>5.9</td>
<td>62.7</td>
</tr>
<tr>
<td>Concerned about interests of third parties</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trainee</td>
<td>70</td>
<td>7.1</td>
<td>8.6</td>
<td>12.9</td>
<td>21.4</td>
<td>50.0</td>
</tr>
<tr>
<td>Trainer</td>
<td>102</td>
<td>6.9</td>
<td>9.8</td>
<td>13.7</td>
<td>11.8</td>
<td>57.8</td>
</tr>
</tbody>
</table>

**SPECIFIC SITUATIONS**

For most specific situations participants reported using ICPC codes ‘often’ or ‘always’ (Table 3). However, repeat prescriptions were seldom or never coded by most participants. Medical examinations were performed by a minority of the trainees and less than half of the trainers and were reported ‘never’ being given ICPC codes by over half of the participants. The most frequent situations or reasons for non-coding mentioned in the free text were: too busy / little time (n=15), telephone consultation (n=13), problem already coded earlier (n=13), no proper code available (n=11), cervical smear (n=8).
Table 3. Specific situations in which ICPC coding is (not) applied

<table>
<thead>
<tr>
<th>How often do you use an ICPC code in the following situations?</th>
<th>n</th>
<th>Never</th>
<th>Seldom</th>
<th>Sometimes</th>
<th>Often</th>
<th>Always</th>
<th>Does not occur</th>
</tr>
</thead>
<tbody>
<tr>
<td>Repeat prescriptions (not during a consultation)</td>
<td>Trainee</td>
<td>69</td>
<td>43.5</td>
<td>17.4</td>
<td>10.1</td>
<td>4.3</td>
<td>5.9</td>
</tr>
<tr>
<td></td>
<td>Trainer</td>
<td>101</td>
<td>39.6</td>
<td>20.8</td>
<td>7.9</td>
<td>6.9</td>
<td>8.9</td>
</tr>
<tr>
<td>Medical examinations (mortgage / driving license / diving etc)</td>
<td>Trainee</td>
<td>69</td>
<td>17.4</td>
<td>0.0</td>
<td>1.4</td>
<td>2.9</td>
<td>5.8</td>
</tr>
<tr>
<td></td>
<td>Trainer</td>
<td>99</td>
<td>28.3</td>
<td>3.0</td>
<td>3.0</td>
<td>4.0</td>
<td>7.1</td>
</tr>
<tr>
<td>Complex problems</td>
<td>Trainee</td>
<td>71</td>
<td>1.0</td>
<td>1.0</td>
<td>15.5</td>
<td>39.8</td>
<td>43.7</td>
</tr>
<tr>
<td></td>
<td>Trainer</td>
<td>103</td>
<td>3.9</td>
<td>4.9</td>
<td>16.7</td>
<td>30.4</td>
<td>40.2</td>
</tr>
<tr>
<td>Extensive differential diagnosis</td>
<td>Trainee</td>
<td>71</td>
<td>2.8</td>
<td>4.2</td>
<td>4.2</td>
<td>29.6</td>
<td>59.2</td>
</tr>
<tr>
<td></td>
<td>Trainer</td>
<td>102</td>
<td>1.0</td>
<td>1.0</td>
<td>13.6</td>
<td>30.1</td>
<td>40.2</td>
</tr>
<tr>
<td>Rare diseases</td>
<td>Trainee</td>
<td>71</td>
<td>-</td>
<td>-</td>
<td>2.8</td>
<td>21.1</td>
<td>76.1</td>
</tr>
<tr>
<td></td>
<td>Trainer</td>
<td>102</td>
<td>1.0</td>
<td>2.9</td>
<td>5.9</td>
<td>25.5</td>
<td>64.7</td>
</tr>
<tr>
<td>Emergencies</td>
<td>Trainee</td>
<td>71</td>
<td>-</td>
<td>-</td>
<td>7.0</td>
<td>29.2</td>
<td>64.8</td>
</tr>
<tr>
<td></td>
<td>Trainer</td>
<td>96</td>
<td>5.2</td>
<td>6.3</td>
<td>35.4</td>
<td>52.1</td>
<td>1.0</td>
</tr>
</tbody>
</table>

**RELATIONSHIP BETWEEN BARRIERS AND CODING**  
The correlation between the self-reported coding performance (scale 1-5, see Table 1) and the actual coding percentage was 0.49 for trainees and 0.81 for trainers (both p<0.01). Correlations (Table 4) between most trainers’ motivation-related or ICPC-related barriers and self-reported or actual coding were moderate to high. Generally the correlations of the trainees were lower than those for trainers. ‘ICPC coding is unpleasant to use’ correlated highest with self-reported and actual coding, both for trainers and trainees. Correlations between external factors and self-reported / actual coding were low and mostly non-significant.

**MULTIVARIATE REGRESSION**  
The only barrier that predicted trainees’ (n=65) self-reported coding performance significantly was ‘ICPC coding is unpleasant to use’ (beta =-0.27), with 7% variance explained. For trainers (n=96), two predictors were significant: ICPC coding is unpleasant to use (beta =-0.30) and takes too much time (beta =-0.31), variance explained was 26%.
### Table 4. Pearson correlation between the barriers of trainees and trainers and their self-reported coding performance and actual coding percentage

<table>
<thead>
<tr>
<th>Barriers</th>
<th>n</th>
<th>r self-reported coding performance</th>
<th>r actual coding percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Motivation</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Takes too much time</td>
<td>Trainee 68/69</td>
<td>-0.30**</td>
<td>-0.26*</td>
</tr>
<tr>
<td></td>
<td>Trainer 99/98</td>
<td>-0.44**</td>
<td>-0.44**</td>
</tr>
<tr>
<td>No personal gain</td>
<td>Trainee 68/69</td>
<td>-0.09</td>
<td>-0.29*</td>
</tr>
<tr>
<td></td>
<td>Trainer 98/97</td>
<td>-0.32**</td>
<td>-0.53**</td>
</tr>
<tr>
<td>No affinity with coding for research purposes</td>
<td>Trainee 68/69</td>
<td>-0.07</td>
<td>-0.24*</td>
</tr>
<tr>
<td></td>
<td>Trainer 98/96</td>
<td>-0.24*</td>
<td>-0.37**</td>
</tr>
<tr>
<td><strong>ICPC Related</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICPC coding is unpleasant to use</td>
<td>Trainee 67/68</td>
<td>-0.33**</td>
<td>-0.35**</td>
</tr>
<tr>
<td></td>
<td>Trainer 98/98</td>
<td>-0.44**</td>
<td>-0.54**</td>
</tr>
<tr>
<td>ICPC Coding is difficult</td>
<td>Trainee 67/68</td>
<td>-0.24*</td>
<td>-0.21</td>
</tr>
<tr>
<td></td>
<td>Trainer 99/98</td>
<td>-0.30**</td>
<td>-0.35**</td>
</tr>
<tr>
<td>ICPC is not sufficiently refined</td>
<td>Trainee 67/68</td>
<td>-0.15</td>
<td>-0.27*</td>
</tr>
<tr>
<td></td>
<td>Trainer 99/98</td>
<td>-0.14</td>
<td>-0.25*</td>
</tr>
<tr>
<td><strong>External</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Should be better paid</td>
<td>Trainee 67/68</td>
<td>-0.03</td>
<td>-0.17</td>
</tr>
<tr>
<td></td>
<td>Trainer 99/98</td>
<td>-0.19</td>
<td>-0.16</td>
</tr>
<tr>
<td>Concerned about interests of third parties</td>
<td>Trainee 68/69</td>
<td>-0.11</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>Trainer 99/98</td>
<td>-0.21*</td>
<td>-0.20</td>
</tr>
</tbody>
</table>

* p<0.01  
** p<0.05  
^a^ 5 point Likert scale

For trainees (n=66) the only significant predictor for the actual coding percentage was ‘ICPC coding is unpleasant to use’ (beta=-0.32); with 10% explained variance. For trainers (n=95), three predictors were significant: ICPC coding is unpleasant to use (beta=-0.31), no personal gain (beta=0.32) and coding is difficult (beta=-0.22); variance explained was 41%.
Discussion

MAIN FINDINGS
We surveyed GP-trainees’ and trainers’ barriers towards ICPC coding and compared the results with their extracted EPR data. Coding percentages were high and most participants experience no or limited coding barriers. The lack of refinement of the ICPC system seems a clear exception to this. Furthermore, almost 30% of trainees reported difficulty with coding as a barrier and a quarter of the participants thought time was a barrier. Many trainers and trainees reported less frequent ICPC coding of repeat prescriptions (Table 3) and administrative actions (Table 1). These categories, however, were excluded when calculating coding percentages and therefore did not influence these scores. ICPC coding was also reported as less frequent for telephone consultations, as confirmed by the EPR extraction data where telephone consultations accounted for about one third of the non-ICPC-coded contacts (data not shown). The proportion of telephone consultations performed by trainees and trainers is lower (12% trainees, 17% trainers), so telephone consultations are likely to be overrepresented in the non-coded contacts.27

No specific situations in which ICPC codes are structurally not applied were identified, either in the questionnaire items (Table 3) or in the free text, suggesting that EPR-extracted ICPC data are not biased other than by non-coded telephone consultations. The trainers’ estimation of their coding performance is fairly accurate, correlating strongly with the actual coding percentage. For trainees this correlation was lower. Trainers’ ICPC coding behaviour seems somewhat more influenced by perceived barriers than the trainees’ coding behaviour.

Motivational and ICPC-related barriers are related to coding performance, unlike external factors (concern about interests of third parties, wish for better payment). For both groups the extent that ICPC coding is experienced as unpleasant to use is predictive for their behaviour. For the trainers the difficulty and the lack of personal gain are also important factors regarding their coding behaviour.

COMPARISON TO OTHER LITERATURE
The coding percentages we found match Jordan et al. in a review of 24 studies.18 De Lusignan suggested that feedback improves coding performance.31 Our results support this; participants received feedback about their coding percentage after each data extraction and the mean coding percentages improved over time. Our findings
also confirm the criticism about the incompleteness of ICPC described by Botsis et al.\textsuperscript{24} Based on an extended Norwegian version of ICPC-2, the authors conclude that the reduction into crude classes ignores the complexity of clinical problems and is therefore inappropriate for clinical work. Almost half of the trainees and a third of the trainers in our study agreed to some extent that ICPC suffers from insufficient refinement. Increasing the diagnostic categories has disadvantages, as stated by Tai et al.\textsuperscript{32} The differences between the ‘picking lists’ of four EPR systems resulted in more diversity rather than consistency. To overcome this, the authors plead (unlike Botsis et al.) for a limited list of codes.

De Lusignan found that ‘primary care clinicians think that definite diagnosis is often an anathema in primary care and that it can stigmatise patients or damage relationship’\textsuperscript{19} However, the ICPC system contains enough symptom codes for users to apply non-definite codes, although difficulties finding the proper code were acknowledged by over half the participants surveyed. The second issue, stigmatising patients (e.g. patients who are unwilling to be diagnosed with a ‘personality disorder’ or ‘depression’), was not directly requested, but never mentioned in free text. About 15\% of the participants (partly) agreed to having concerns about the interests of third parties, like health insurance companies. No correlation appeared between this barrier and the actual coding percentage, and only a weak correlation between this barrier and self-reported performance of the trainers. These findings, together with the high coding percentages encountered, strongly suggest that fear of stigmatising or of damaging the relationship with patients was not relevant in our sample.

\textbf{LIMITATIONS}

When selecting participants for the main study, two GP trainers who did not use or did not want to use ICPC coding were excluded\textsuperscript{27} and another 14 trainers declined to participate. Therefore, this study may suffer from underreporting of coding barriers. Furthermore, all trainers and trainees were allied to a single institute which could influence the representativeness of our sample. Representativeness may also be impaired due to the inclusion of GPs working in training practices only.

The free text response revealed that not all participants were familiar with or agreed to the guidelines by the Dutch College of General Practitioners (NHG), that every patient encounter must be coded, as demonstrated by the thirteen participants who responded that they did not code ‘if the problem had already been coded earlier’. The specific situations prone to non-coding were surveyed by self-report
only. They were not objectified by free text analyses of the EPR, or by video analyses or otherwise.

**IMPLICATIONS AND FUTURE RESEARCH**

Indications of bias found were limited and restricted, making the use of EPR data for research purposes and for patient monitoring in the GP specialty training a reasonably valid method, preferable to hand-written or separate electronic logbooks. However, there is room for improvement. To further reduce the chance of biased results, the content of the non-coded telephone consultations should be addressed. Furthermore, the results of the regression highlight the importance of increasing motivation for coding and making ICPC coding more pleasant to use. Future research should evaluate interventions directed at increasing coding and explaining its benefits, by providing feedback about contact numbers, diagnoses, prescriptions, referrals and coding percentages compared to adjacent practices and national reference numbers. ICT innovations providing the user with likeable functionality and better patient care such as care management reminders and computer assisted diagnosis, which often depend on proper ICPC coding, may increase motivation.

ICPC related barriers can be overcome by easier ICPC registration using user-friendly technical solutions.

The effect of making it impossible to finalize encounters without providing an ICPC code may also be studied, although this can prove annoying to users. One trainee-trainer couple using such a system (for all encounters) achieved coding percentages of 100%. Future research should also address the content of the non-coded telephone consultations.

**CONCLUSION**

This study gave insight into the coding behaviour of both GP trainers and trainees and into the resulting quality of the EPR data. Results showed that coding percentages were high. Further improvement could be realized by increasing motivation and by making the ICPC more pleasant to use. Since no other specific areas prone to non-coding were found, EPR derived data seem to be biased mainly by non-coded telephone consultations. Besides this, EPR data seem to provide a true reflection of the patients seen by trainers and trainees, making these data suitable for research and educational purposes.
REFERENCES

2. van der Linden MW, Westert GP, de Bakker DH, Schellevis FG. Second National Study into Diseases and Actions in General Practice [Tweede Nationale Studie naar ziekten en verrichtingen in de huisartspraktijk], 2004. Utrecht (Bilthoven): Dutch Institute for Health Services Research (NIVEL) and National Institute of Public Health and Environment (RIVM).
16. Britt H, Angelis M, Harris E. The reliability


30. de Jong J, Visser MR, Wieringa-de Waard M. Steering patient mix is not as easy as it seems. Results of a randomised controlled trial. Submitted for publication.


Opening the black box:
the patient mix of
GP trainees

Jip de Jong, Mechteld R.M. Visser, Jacob Mohrs, Margreet Wieringa-de Waard

Br J Gen Pract 2011; e650-7; DOI: 10.3399/bjgp11X601361
Abstract

**Background** The variety of health problems (patient mix) that medical trainees encounter is presumed to be sufficient to master the required competencies.

**Aim** To describe the patient mix of GP trainees, to study differences in patient mix between first-year and third-year GP trainees, and to investigate differences in exposure to sex-specific diseases between male and female trainees.

**Design and setting** Prospective cohort study in Dutch primary care.

**Method** During a 6-month period, aggregated data about International Classification of Primary Care diagnosis codes, and data on the sex and age of all contacts were collected from the electronic patient record (EPR) system.

**Results** Seventy-three trainees participated in this study. The mean coding percentage was 86% and the mean number of face-to-face consultations per trimester was 450.0 in the first year and 485.4 in the third year, indicating greater variance in the number of patient contacts among third-year trainees. Diseases seen most frequently were: musculoskeletal (mean per trimester = 89.2 in the first year/91.0 in the third year), respiratory (98.2/92.7) and skin diseases (89.5/96.0). Least often seen were diseases of the blood and blood-forming organs (5.3/7.2), male genital disorders (6.1/7.1), and social problems (4.3/4.2). The mean number of chronic diseases seen per trimester was 48.0 for first-year trainees and 62.4 for third-year trainees. Female trainees saw an average of 39.8 female conditions per trimester — twice as many as male trainees (mean = 21.3).

**Conclusion** Considerable variation exists in the number of patient contacts. Differences in patient mix between first- and third-year trainees seem at least partly related to year-specific learning objectives. The use of an EPR-derived educational instrument provides insight into the trainees’ patient mix at both the group and the individual level. This offers opportunities for GP trainers, trainees, and curriculum designers to optimise learning when exposure may be low.
Introduction

In most contemporary medical-specialty curricula, trainees acquire the demanded competencies through work-based learning,¹ which allows competency development through clinical exposure to real patients (patient mix). Earlier research suggested that the patient mix of GP trainees differs from that of their trainers; trainees were exposed to more minor ailments and fewer chronic diseases and severe conditions.²–⁷ These studies were, however, small scale, covered a small period, or were relatively old.²;⁷–¹²

An explanation for disparities between the patient mix of trainers and trainees, suggested by Stubbings and Gowers, may be that the trainer is the senior, more experienced partner, thereby attracting older and more complex patients.⁷ The assigning behaviour of medical receptionists was not found to explain such disparities.¹³

Low exposure in some fields of patient mix was also found in undergraduate medical curricula, clerkships, and other specialty training programmes.⁵;¹⁴–²² Ericsson’s ‘Deliberate practice’ theory argues that low exposure has a negative impact on competency development. It assumes that becoming an expert requires sufficient practical experience and appropriate reflection, which can be stimulated by feedback from coaches or trainers.²³ Within this framework, patient mix is an important training condition. Empirical evidence supports the importance of adequate patient mix to the effectiveness of rotations²⁴ and the instructional quality of GP settings.²⁵ The importance of an appropriate patient mix was also acknowledged by the World Federation For Medical Education (WFME), a strategic partner of the World Health Organization (WHO). In the WFME global standards for quality improvement in postgraduate education, it is stated: ‘The training must expose the trainee to a broad range of experience in the chosen field of medicine.’²⁶

Electronic data extracted from electronic patient record (EPR) systems have proven valid for providing insight into the patient mix confronting GP trainees.⁹;²⁷–²⁹ In the Netherlands, most GPs keep detailed EPRs, including a standardised system of diagnosis codes: the International Classification of Primary Care (ICPC).³⁰ Stimulated by the Dutch College of General Practitioners, all EPR systems now include the ICPC-1 codes. This allows uniformity in descriptions of the patient mix of GP trainees. The WHO has accepted ICPC as a reason for encounter classification, and users may use it as a classification for primary care or general practice wherever applicable.
Detailed descriptions of patient mix can be used to determine whether, and in which areas, low exposure actually exists. Teaching and rotation programmes may be adapted accordingly. For instance, trainees confronted with many older patients during their first-year practice period could benefit from subsequent assignment to a practice with many young families, or a rotation in a child health centre.

For GP trainers, detailed knowledge about the actual patient mix of their trainees is necessary to optimise this mix for specific learning objectives. The first-year curriculum focuses on minor ailments and daily care, while chronic-disease management is a formal learning objective in the third year. GP trainers can tailor their trainees’ patient mixes to these objectives.

Several authors have mentioned differences between male and female physicians regarding the number of contacts with patients with female conditions.\textsuperscript{7;31–33} Female trainees saw more female conditions than male trainees did, leaving the latter with relatively low exposure. The opposite could apply to male conditions, causing relatively low exposure for female trainees.

DEFINITION OF PATIENT MIX
For this study, patient mix was defined as the quantity and variety of diseases and the age and sex distribution of patients. This broad definition was used, as patient demographic characteristics can be relevant for learning objectives.\textsuperscript{4;5;24;25}

RESEARCH QUESTIONS
The research questions were as follows:
• what patient mix do GP trainees encounter?
• how does patient mix differ between first-year and third-year trainees regarding number of patient contacts, mean patient age, and number of diagnosis codes within the different organ systems?
• how do male and female trainees differ in their exposure to sex-specific diseases?
• how many patients with chronic diseases do trainees encounter?
Method

PARTICIPANTS
Participants were 49 first-year and 24 third-year GP trainees at the Institute of GP Specialty Training of the Academic Medical Center University Amsterdam.

SETTING
This study was conducted in 2008–2009 in GP practices affiliated with the Institute of GP Specialty Training, which facilitates a 3-year training programme in which first- and third-year trainees are stationed in GP training practices. In their second year, trainees do clinical rotations. A full-time working week was 36 hours, including 6.5 hours of modular education at the institute and 29.5 hours in the practices.

DESIGN AND PROCEDURES
This was a prospective cohort study. Between March and December 2008, GP trainers with trainees in their practices starting a training period of 9–12 months were approached for inclusion. For each trainee, data were extracted for 6 months. Data-extraction software was developed for four different types of EPR, covering approximately 80% of the EPRs used by the affiliated training practices.

First, the trainers were asked to participate. The trainees of those who agreed were invited for the study. All GP trainers and trainees were asked for informed consent and provided with reference cards detailing the basic ICPC coding rules and some standard coding problems and solutions. All trainees attended a session that introduced the study and explained the content of the reference card. The regular training programme also addresses ICPC coding, and ICPC-coding assistance was available by email or telephone throughout the study.

DATA EXTRACTION AND INSTRUMENT
Data were extracted directly from the EPR every 3 months by either the GPs or the researchers. The raw extraction data were processed by multiple software routines using MS Visual Basic routines in MS Access. Decision rules were programmed to ensure logic and consistency in counting the codes (Text box 1). The resulting aggregate data were emailed to the researchers. No information about individual patients could be recovered. The data contained information about ICPC diagnosis codes (I) and patient contacts (II). The latter were aggregated according to sex, age, and consultation type (telephone, face-to-face or home visit).
Text box I. Decision rules

ICPC diagnosis codes (I)

- Each newly entered ‘distinct’ diagnosis code was counted. In case of multiple diagnosis codes, all were counted, even if they took place in the same consultation.
- Codes for specific patients on specific days that were entered twice or more were counted only once.
- All codes were counted for patients who were seen on separate days for similar problems, leading to the same ICPC codes. (Note that this approach is different from calculating morbidity rates).³⁵
- For diagnosis codes linked to episodes of care, the episode code was counted only if no other new diagnosis code was entered during that contact; the episode code was ignored if a distinct diagnosis code was entered and linked to the episode code (for example, if pneumonia was linked to a chronic obstructive pulmonary disease (COPD) episode, only the pneumonia code was counted, whereas COPD would have been counted in the absence of the pneumonia code during that specific consultation or visit).
- If no diagnosis code (distinct or episode) was present, the consult was counted as non-coded.
- Diagnosis codes linked to administrative tasks or repeat prescriptions were ignored.

Contact-type codes (II)

- Contact codes were categorised as face-to-face, telephone, or home visit.
ICPC DIAGNOSIS CODES
Developed by the ICPC Working Party and published in 1987 by the World Organization of National Colleges, Academies (WONCA), the ICPC consists of 17 chapters covering the organ systems, a general chapter, and a social chapter. WHO includes the ICPC (Version 2) within its FIC (Family of International Classifications). The differences between ICPC-1 and ICPC-2 regarding diagnosis codes are minor. The ICPC coding system allows the coding of both symptom diagnoses (for example, headache, ankle complaints) and ‘definite’ diagnoses (for example, pneumonia, appendicitis, hypothyroidism).

CHRONIC DISEASES
The list of chronic diseases published by Knottnerus and colleagues was used to count the number of chronic diseases the trainees encountered.

DATA ANALYSIS
Descriptive analyses were used to report the mean number of contacts per trimester, encounters per ICPC chapter, and coding percentage. Differences between male and female trainees and between first- and third-year trainees were tested using analysis of variance and t-tests; 95% significance level.

The coding percentage was calculated by dividing the total number of ICPC diagnosis codes in a trimester by that same number plus the number of non-coded contacts. Because there can be multiple diagnosis codes per contact, this formula could overestimate the coding percentage. The coding percentages were therefore corrected by multiplying the non-coded contacts in each case by the individual mean number of ICPC diagnosis codes per contact.

Data from part-time trainees were adjusted to give values as if they worked full-time, by multiplying the absolute numbers of diagnoses and contact codes by 1/(part-time percentage/100).
Results

PARTICIPANTS
During the inclusion period, trainees were starting training periods in 109 practices affiliated with the researchers’ institute. The 98 practices that had suitable EPR systems were contacted. The sample included 73 GP training practices, 49 first-year trainees, and 24 third-year trainees. Reasons for non-inclusion (n=25) are displayed in Figure 1. Eleven (15%) practices were single-handed, 17 (23%) were dual-handed, and 45 (62%) were group practices.

The mean age of trainees was 31.2 years (standard deviation [SD]=3.2 years): 30.2 years (SD=2.5) for first-year trainees and 33.2 years (SD=3.7) for third-year trainees. Six first-year trainees (12%) and eight third-year trainees (33%) worked part-time, with a mean part-time percentage of 86% corresponding to 31 hours per week.

Figure 1: Flow chart

109 screened for participation

98 Trainers asked to participate

73 Trainees included
49 Year 1 (14 male, 35 female)
24 Year 3 (5 male, 19 female)

3 (n=8) or 6 (n=65) months basic registration completed
49 Year 1
24 Year 3

II EPR unsuitable

2 Not used to diagnosis coding
5 A recent or forthcoming switch to a different (unsuitable) EPR
2 Illness/deceased
2 No GP trainee assigned at time of the study
I4 Refusal (mostly due to time constraints)

EPR = electronic patient record
CODING PERCENTAGE
The mean (corrected) coding percentage of 86% (SD=13.2) was similar in both training years (86% versus 87%). Data regarding the ICPC codes for one trainee with a very low coding percentage (27%) were disregarded.

CONSULTATIONS AND HOME VISITS
The mean number of contacts per trimester was 583.8. Table 1 presents the mean number of contacts in the first and third training years, averaged over two trimesters. There were no significant differences between the first and third years, although the SD for face-to-face consultations in year 3 was twice as high as for year 1, indicating greater variance in the number of patient contacts among third-year trainees.

The mean number of ICPC diagnosis codes per trimester, and the mean number of non-coded contacts are also presented in Table 1. The mean number of ICPC codes per contact was 1.13 in the first year and 1.11 in the third year. The total number of problems encountered can thus be estimated at 621.0 in the first year and 661.9 in the third year.

Table 1.
Mean contact frequencies, number of ICPC diagnosis codes and number of non-coded contacts for first and third-year trainees, averaged over two trimesters

<table>
<thead>
<tr>
<th></th>
<th>First-year trainees</th>
<th>Third-year trainees</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N=49</td>
<td>N=24</td>
</tr>
<tr>
<td><strong>Mean per trimester (SD)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Face-to-face</td>
<td>450.0 (86.6)</td>
<td>485.4 (159.1)</td>
</tr>
<tr>
<td>Telephone</td>
<td>64.0 (35.9)</td>
<td>75.4 (56.6)</td>
</tr>
<tr>
<td>Home visit</td>
<td>35.6 (30.2)</td>
<td>35.5 (26.7)</td>
</tr>
<tr>
<td>Total contact codes</td>
<td>549.6 (111.3)</td>
<td>596.3 (181.1)</td>
</tr>
<tr>
<td>ICPC diagnosis codes*</td>
<td>537.6 (153.1)</td>
<td>576.4 (201.7)</td>
</tr>
<tr>
<td>Non-coded contacts**</td>
<td>75.9 (71.8)</td>
<td>76.8 (86.8)</td>
</tr>
<tr>
<td></td>
<td>251 / 641</td>
<td>252 / 797</td>
</tr>
<tr>
<td></td>
<td>5 / 139</td>
<td>11 / 224</td>
</tr>
<tr>
<td></td>
<td>0 / 147</td>
<td>6 / 120</td>
</tr>
<tr>
<td></td>
<td>302 / 782</td>
<td>301 / 939</td>
</tr>
<tr>
<td></td>
<td>337 / 884</td>
<td>261 / 955</td>
</tr>
<tr>
<td></td>
<td>0 / 392</td>
<td>4 / 372</td>
</tr>
</tbody>
</table>

* Number of ICPC codes, averaged over two trimesters, corrected for part-time work
** Mean number of non-coded contacts, averaged over two trimesters, corrected for part-time work
AGE AND SEX DISTRIBUTION OF PATIENTS

Figure 2 presents the age distribution of the patients. Third-year trainees saw significantly more patients aged between 45 and 64 years than did first-year trainees ($t(71) = -2.58, p = 0.012$). No significant differences between first-year and third-year trainees were found for the other age groups.

Female patients were seen significantly more (mean=344.6, SD=92.3) than male patients (mean=220.3, SD=59.3) ($t(144) = -9.68, p = 0.00$). No differences were found between male and female trainees regarding the sex distribution of patients ($t(71) = 0.503, p = 0.61$).

Figure 2. Means and standard deviation for number of contacts per age-group for first and third-year trainees, averaged over two trimesters.
ORGAN SYSTEMS

Figure 3 presents the patient mix by organ system and the mean numbers of diagnosis codes per trimester. Musculoskeletal (year 1 mean per trimester [SD]/year 3 mean [SD] = 89.2 [28.6] / 91.0 [39.2]), respiratory (98.2 [30.8] / 92.7 [34.7]), and skin diseases (89.5 [26.3] / 96.0 [32.8]) were seen most often. Diseases of the blood and blood-forming organs (5.3 [3.3] / 7.2 [4.8]), male genital disorders (6.1 [3.4] / 7.1 [6.6]), and social problems (4.3 [3.2] / 4.2 [3.1]) were seen the least. Significant differences were found between years 1 and 3 for blood and blood-forming diseases, (t(70) = -2.02, p = 0.047), psychiatric diseases (t(70) = -2.56, p = 0.013), and metabolic diseases (t(70) = -2.53, p = 0.014).

Figure 3. Means and standard deviations of ICPC codes for first- and third-year trainees, averaged over two trimesters

CONDITIONS SEEN MOST FREQUENTLY

Table 2 presents the 20 conditions most often seen in either training year. Differences between first- and third-year trainees were minor: otitis media was in the top 20 of first-year trainees (third-year rank 23, mean = 4.9 contacts per trimester), whereas diabetes mellitus was in the top 20 of the third-year trainees (first-year rank 26, mean = 4.5 contacts per trimester).
Table 2. Mean contacts per trimester for the conditions seen most frequently in Years 1 and 3, averaged over two trimesters

<table>
<thead>
<tr>
<th>Rank</th>
<th>Year 1</th>
<th>Mean</th>
<th>Year 3</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Acute Respiratory Infection</td>
<td>28.7</td>
<td>Acute Respiratory Infection</td>
<td>24.0</td>
</tr>
<tr>
<td>2</td>
<td>Cough</td>
<td>15.8</td>
<td>Hypertension</td>
<td>14.2</td>
</tr>
<tr>
<td>3</td>
<td>Hypertension</td>
<td>11.3</td>
<td>Cough</td>
<td>12.8</td>
</tr>
<tr>
<td>4</td>
<td>Dermatomycosis</td>
<td>9.6</td>
<td>Dermatomycosis</td>
<td>10.9</td>
</tr>
<tr>
<td>5</td>
<td>Otitis media</td>
<td>8.5</td>
<td>Low Back Pain</td>
<td>8.5</td>
</tr>
<tr>
<td>6</td>
<td>Low Back Pain</td>
<td>7.8</td>
<td>Cystitis or Other Urinary Infection</td>
<td>8.3</td>
</tr>
<tr>
<td>7</td>
<td>Cystitis or other Urinary Infection</td>
<td>7.5</td>
<td>Acute or Chronic Sinusitis</td>
<td>7.8</td>
</tr>
<tr>
<td>8</td>
<td>Acute or Chronic Sinusitis</td>
<td>7.5</td>
<td>Other Musculoskeletal Disease</td>
<td>7.4</td>
</tr>
<tr>
<td>9</td>
<td>Weakness/Tiredness</td>
<td>7.4</td>
<td>Contact / Allergic Dermatitis</td>
<td>7.1</td>
</tr>
<tr>
<td>10</td>
<td>Acute Bronchitis or Bronchiolitis</td>
<td>7.0</td>
<td>Weakness or Tiredness</td>
<td>7.1</td>
</tr>
<tr>
<td>11</td>
<td>Knee Symptom/Complaint</td>
<td>6.6</td>
<td>Diabetes Mellitus</td>
<td>6.8</td>
</tr>
<tr>
<td>12</td>
<td>Dermatitis or Atopic Eczema</td>
<td>6.5</td>
<td>Acute Bronchitis or Bronchiolitis</td>
<td>6.6</td>
</tr>
<tr>
<td>13</td>
<td>Contact or Allergic Dermatitis</td>
<td>6.3</td>
<td>Abdominal Pain/Cramps (General)</td>
<td>6.4</td>
</tr>
<tr>
<td>14</td>
<td>Localized Abdominal Pain</td>
<td>6.2</td>
<td>Musculoskeletal Injury</td>
<td>6.4</td>
</tr>
<tr>
<td>15</td>
<td>Shoulder Symptoms</td>
<td>6.0</td>
<td>Asthma</td>
<td>6.1</td>
</tr>
<tr>
<td>16</td>
<td>Otitis Externa</td>
<td>5.9</td>
<td>Knee Symptom/Complaint</td>
<td>5.9</td>
</tr>
<tr>
<td>17</td>
<td>Asthma</td>
<td>5.9</td>
<td>Dermatitis/Atopic Eczema</td>
<td>5.7</td>
</tr>
<tr>
<td>18</td>
<td>Gastroenteritis</td>
<td>5.4</td>
<td>Constipation</td>
<td>5.7</td>
</tr>
<tr>
<td>19</td>
<td>Musculoskeletal Injury</td>
<td>5.4</td>
<td>Conjunctivitis</td>
<td>5.3</td>
</tr>
<tr>
<td>20</td>
<td>Pneumonia</td>
<td>5.3</td>
<td>Gastroenteritis</td>
<td>5.3</td>
</tr>
</tbody>
</table>

**CHRONIC DISEASES**

The mean number of chronic diseases seen per trimester was 48.0 (SD=21.9; 39.5 first trimester, 56.5 second trimester) for first-year trainees, and 62.4 (SD=32.6; 55.3 first trimester, 78.3 second trimester) for third-year trainees. The seven chronic conditions seen most frequently were the same for both years: diabetes mellitus, asthma, constitutional eczema, emphysema or chronic obstructive pulmonary disease (COPD), depression, heart failure, and leg ulcer.

**SEX-SPECIFIC CONDITIONS**

Female trainees saw almost twice as many (mean per trimester=39.8, SD=24.1) female conditions as did male trainees (mean=21.3, SD=9.2). This difference was significant (F(1,68)=7.38, p=0.008), irrespective of training year. For male conditions, no similar pattern was found.
Discussion

SUMMARY
The patient mix of 73 GP trainees was studied by using EPR-derived data extractions, and compared the first and third training years. It was found that the patient mix is generally well spread out over all ages and organ systems, although it is possible to indicate areas of relative low exposure. Some differences were found between the patient mixes of first- and third-year trainees that seemed related to year-specific learning objectives. Third year trainees saw more chronic diseases than first years did. It was also found that female trainees saw more female conditions than their male counterparts.

Contrary to expectations, no higher contact frequencies were observed in the third training year, although frequencies varied more than in the first year. Closer examination showed that this was largely due to a number of trainees who saw exceptionally large numbers of patients. This kind of information about the patient mix of individual trainees was not known before monitoring was started. The results presented here may in general seem satisfactory for the group as a whole, but since the variance in contacts and exposure to the different organ systems was substantial, individual trainees may, perhaps unknowingly, be suffering from low exposure in certain areas. This study’s findings show the importance of monitoring the patient mix for each individual GP trainee.

STRENGTHS AND LIMITATIONS
Using EPR data to describe patient mix has been successful before.9,14,35 One strength of the present study was the use of data from EPR systems combined with decision rules to describe patient mix. Many earlier studies used handwritten or digital logbooks intended solely for educational purposes.5,14,36 These type of data are subject to social-desirability bias by the trainee. Extracted from EPR systems, the data of the present study were not biased by any threat of formal assessment. Such systems can be of great educational value.

Diagnosis-coding percentages were satisfying, at about 86%, although it is not possible to be sure whether this patient mix was representative. For example, the data could have been biased if trainees tended not to code complex cases.

Although ICPC coding is useful for describing patient mix, its validity is potentially diminished by its dependence on the diagnostic competence of the train-
ees and the adequacy of the attribution of the codes (for example, definite instead of symptom diagnosis).37

Although all trainees were affiliated with the same specialty-training institute, the authors do not think that this influenced the results, as other studies in different times, scales, and areas have shown similar results.

COMPARISON WITH EXISTING LITERATURE
The trainees were exposed to diseases of all organ systems. Most prominent were respiratory, skin, and musculoskeletal diseases, followed by digestive, circulatory and ear diseases. This exposure pattern reflects the general morbidity pattern in the Netherlands (2nd Dutch National Survey of General Practice [DNSGP-2]).37 These findings seem robust, as other studies (for example, the 1991 and 1992 British Morbidity Surveys) have found similar exposure patterns, although these studies were small scale, covered a small period, or were relatively old.2;7–12

Trainees were less likely to encounter diseases of the blood and blood-forming organs, male genital disorders, and social problems. This was also the case in the DNSGP-2 and the British Morbidity Survey (excepting social problems), meaning that licensed GPs are also less likely to encounter these conditions. From an educational perspective, however, potential underexposure in these areas is conceivable. Several authors have warned about underexposure to chronic conditions.38–41 Earlier studies report that GP trainees saw fewer chronic conditions than did British principals4 or Dutch trainers.6 In 2004, Darer and colleagues found that US physicians considered their training in chronic illness care inadequate.39 In the present study, exposure to chronic diseases increased steadily during training. This should be considered positive, as the third-year curriculum emphasises chronic and complex conditions as learning objectives. The actual number of chronic conditions encountered by trainees is probably higher than the findings of this study, as the software was programmed to count actual health problems and not intercurrent chronic conditions. For example, coughing and dyspnoea in an 82-year-old patient with a history of diabetes, stroke, and Parkinson’s disease may be coded simply as a common cold, even if all comorbidity was taken into account.

It was found that female conditions were seen more often by female trainees. This is in line with studies on the clinical skills of medical students and family medicine residents.32;33 Levy and Merchant found that male students received more experience with male-specific examination skills, but the present data do not show
that male GP trainees saw more male-specific conditions. The low number of codes regarding social conditions might be due to under-reporting. Although trainees may be aware of a patient’s social problems, they may not be coded unless they were the complaints presented by the patient (for example, in the interest of time, carefulness, or legal considerations).

Although there are no data available about the trainees’ competence development in areas of low exposure, these areas may need attention by GP trainers and curriculum designers. They should be aware that adequate supervision is needed in areas of low exposure, in order to maximise the learning experiences. Another option for addressing potential lacunas is to extend the training with clinical rotations in these areas.

**IMPLICATIONS FOR RESEARCH**

The higher variation of contact frequencies in the third year, mainly due to several trainees seeing an exceptional large amount of patients, is interesting from an educational perspective. Trainees who see large numbers of patients might build more experience, learn more, and perform better. It would be interesting to identify whether this is due to individual characteristics of these trainees, their trainers or the patient supply within the training practice.

Another question for further study involves whether steering trainees’ patient mix can affect competence development. For example, the results of this study suggest that steering patients with female conditions to male trainees may be rational.

If approached with caution, the results may help to identify specific practices with high concentrations of patients with specific characteristics. Future studies could investigate whether the intentional allocation of trainees with specific learning objectives to suitable practices improves the patient mix.

It may be attractive to formulate minimum quotas or estimates of the number of patients needed to develop competence in certain areas. Although the present results suggest what is normally encountered, caution is advised, as the relation between patient mix and competence development is complex and differs between individuals.
REFERENCES


37. van der Linden M, Schellevis F, Westert G. Morbidity in the population and in general practice: 2nd Dutch National Survey of Gen-


Exploring differences in patient mix in a cohort of GP trainees and their trainers

Jip de Jong, Mechteld R.M. Visser, Margreet Wieringa-de Waard

*BMJ Open* 2011; 1:e000318. doi:10.1136/bmjopen-2011-000318

*Huisarts & Wetenschap* 2012; 55(7): 290-5 (shortened version, in Dutch)
Abstract

**Background** During specialty training for general practice, trainees acquire the required competencies through work-based learning. Previous small-scale and older studies suggest that the patient mix of general practitioner (GP) trainees differs from that of their trainers: trainees are exposed to more minor illnesses, and fewer chronic diseases and severe conditions, which may influence the development of their competency.

**Research question** What are the differences in the patient mix between trainees and trainers?

**Methods** 49 first- and 24 third-year trainees and their trainers (n=114) were included in the study. International Classification of Primary Care (ICPC) contact and diagnosis codes were extracted from electronic patient records over 6 months.

**Results** Trainers had double the number of face-to-face consultations, and treble the number of telephone consultations compared with trainees. The trainees’ patient mix consisted of significantly more patients with eye diseases, ear diseases, respiratory diseases, skin diseases and minor illnesses compared with their trainers. Trainers encountered significantly more patients with circulatory diseases, psychiatric diseases, metabolic diseases, male genital conditions, social problems, and chronic and oncological diseases. Female trainers and trainees encountered almost twice the number of female conditions compared with their male counterparts, while for male conditions, the opposite was found.

**Discussion** Considerable differences between the patient mix of trainers and trainees were found. Specialty trainers and teachers must be aware of areas of low exposure. Trainers should ensure trainees handle more chronic, complex, psychosocial and circulatory conditions.
Introduction

During their specialty training, general practitioner (GP) trainees develop the required competencies through work-based learning involving clinical exposure to an adequate patient mix. The importance of an adequate patient mix was recently confirmed\textsuperscript{1-3} and recognised by several national\textsuperscript{4-6} and international\textsuperscript{7} accreditation standards. The World Federation for Medical Education emphasised in its Global Standards for Quality Improvement for Postgraduate Education\textsuperscript{7} the importance of a broad range of experience in the trainee’s chosen medical field.

But what is an adequate patient mix? The learning curve of individual trainees may vary,\textsuperscript{8;9} and successful work-based learning depends on many factors, such as learning style,\textsuperscript{10} independence\textsuperscript{3} and, especially, supervision.\textsuperscript{3;11} An adequate patient mix should contain enough diverse learning experiences\textsuperscript{12} and should resemble the patient mix the trainees will be confronted with later as licensed GPs.\textsuperscript{13-15} To address this requirement and to determine low exposure areas, the patient mix of trainees and their trainers should be analysed. Previous reports\textsuperscript{14-29} found that the trainees’ patient mix consisted of more minor illnesses and fewer psychosocial, chronic and severe conditions compared with their trainers, demonstrating low exposure to important health conditions.\textsuperscript{30} These studies, however, are relatively old,\textsuperscript{14;16;21-26;29} were small scale (n=8)\textsuperscript{20} or even case studies (n=1),\textsuperscript{14;21;23;26;28;29} and often covered short periods (≤ 4 weeks).\textsuperscript{15;18;22;24} A larger study over a longer period has not been carried out recently.

Trainer insight into trainee patient mix should be encouraged and trainers should aim at a tailored mix to ensure adequate learning experiences.\textsuperscript{21} Tailoring could be achieved, for instance, by instructing the medical receptionist to request patients belonging to specific patient groups to attend the trainee.\textsuperscript{13;21;31}

If the patient mix is deficient due to unavoidable factors (eg, geographical location), a placement in a practice with a complementary patient mix could be arranged by the training institute. A detailed description of the patient mix can be established using logbooks\textsuperscript{32;33} or extracting data from electronic patient records (EPR), as in the present study.\textsuperscript{20;34;35} Because proper medical reporting is considered of paramount importance, EPR systems are valid and reliable, and, unlike logbooks, do not require additional actions. The aim of this study was to investigate the differences in patient mix between GP trainees and their trainers.
Methods

PARTICIPANTS AND STUDY SETTING
The study was conducted between 2008 and 2009 in practices affiliated to the GP specialty training programme of the Academic Medical Center, University of Amsterdam. This training institute facilitates a 3-year course in which first- and third-year trainees work in GP training practices. Trainees are assigned clinical rotations in their second year. Trainees work for 36 h per week, 29.5 h of which are spent in direct patient care. The average GP in the Netherlands worked for 44 h per week in 2001.36

DESIGN AND PROCEDURES
Training practices about to accommodate a new trainee were identified. If trainers agreed to participate, their trainees were also approached for inclusion.

All participants gave informed consent. In the Netherlands, most GPs keep detailed EPRs utilising a standardised system of diagnosis codes set out in the International Classification of Primary Care (ICPC). The ICPC consists of 17 chapters covering the organ systems, a general chapter and a social chapter.

All participants received a visual aid describing the basic ICPC coding rules and some standard problems and solutions. All trainees attended a short session, in which the study was presented and the visual aid was explained. The trainers were instructed individually. GPs not certified as trainers who worked in training practices were included if they supervised the trainee for 1 day/week or more.

Data were extracted over 6 months. Ethics approval was obtained from the Ethical Review Board of the Netherlands Association for Medical Education (NERB-ID 42).

DATA EXTRACTION
We developed software to extract data from the EPR systems.37 The data contained aggregated information about ICPC diagnosis codes including information about the frequency of patient contacts by age group and consultation-type code. The chance of missing a consultation-type code was small since these are used for invoicing. Most EPR systems demand a consultation-type code for each contact with a patient and check for missing codes. The coding percentage was therefore always 100%. Consultation-type codes determined whether the ICPC diagnosis codes were
counted or not. Administration acts or repeat prescriptions were disregarded. The consultation-type codes were translated into three basic contact codes: telephone consultations, face-to-face consultations or home visits. Entering an ICPC diagnosis code in the EPR after seeing a patient is usually not mandatory, so the ICPC diagnosis coding percentage could be below 100%.

**STATISTICAL ANALYSES**

Trainer-trainee differences were tested using ANOVA in SPSS V.18.0. The data of part-time workers were corrected to reflect full-time work. For the patient mix description, we used the mean percentages of the ICPC codes. To calculate these, the non-coded contacts were disregarded for each subject, and the total of the remaining codes was set to 100%.

The coding percentages were calculated by dividing the total number of ICPC diagnosis codes in a trimester by that same number plus the number of non-coded contacts, corrected by the individual mean number of ICPC diagnosis codes per contact. If the coding percentage was below 50%, the subject was excluded from the ICPC code analyses. We analysed several specific diseases and diseases clusters of interest for training purposes. To evaluate the number of chronic and oncological diseases encountered, we used the clusters published by Knottnerus et al.38 For the description of minor illnesses and acute diseases, we asked five experienced GPs to score the appropriateness of (preselected) ICPC codes for these disease categories on a scale of 1-9. The ICPC code was included if the median score was higher than 7 and none of the scores was below 3 (see supplementary appendices 1 and 2).39

To ensure all learning experiences were represented, every ICPC diagnosis code entered during a patient contact was counted and we compared our data with those of the Netherlands Information Network of General Practice (LINH, 2009).40 These latter data were collected from a representative network of 84 general practices throughout the Netherlands, with about 140 GPs.40 The majority were not training practices. There were two differences compared with our data sampling: LINH data are primarily morbidity figures and based on the last ICPC code in a series of contacts (episodes); and the contacts of the medical receptionists and nurse practitioners (NPs) are included in the LINH data. We therefore also compared our results with older data, those of the second Dutch National Survey of General Practice (104 practices) which was carried out in 2001 (DNSGP-2, 2004).41 The advantage of this comparison is that in 2001 only a few NPs were working in GP practices. The DNSGP-2 also used the LINH registration network.
Results

PARTICIPANTS
Ninety-eight GP-training practices were approached and 73 were included in the study. Reasons for non-inclusion (n=25) were: not used to diagnosis coding (n=2), a recent/imminent change to a different (unsuitable) EPR (n=5), illness (n=2), no trainee assigned (n=2) and refusal (n=14).

A total of 49 first-year trainees and 24 third-year trainees were working in these 73 practices. There were 10 single, 17 dual and 46 group practices. There was one trainer in 32 of the training practices, and there were two or more in the remaining 41. The data of 73 trainees and 114 trainers (102 officially certified and 12 supervising) were extracted. The trainees’ mean age was 31.2 (SD 3.24) years and the trainers’ mean age was 50.9 (SD 7.0) years; more than half of the trainers worked full-time (n=59).

CODING PERCENTAGES OF ICPC CONTACT CODES
The mean percentage of the contacts containing at least one ICPC diagnosis code was 86.2% (SD 13.2, range 27.4%-100%) for trainees and 78.9% (SD 22.4, range 2.3%-100%) for trainers. Forty-two trainees and 53 trainers had a coding percentage over 90%. Fourteen trainers and two trainees had a mean ICPC coding percentage below 50% and their ICPC diagnosis code data were disregarded. Of the remaining participants, the mean coding percentage was 88.1% (SD 9.7) for trainees and 85.6% (SD 12.7) for trainers.

CONSULTATIONS AND HOME VISITS
The mean number of trainee consultations per trimester was 461.2 (SD 115.6) for face-to-face contact, 67.7 (SD 43.7) for telephone contact and 35.6 (SD 28.9) for home visits.

Trainers averaged 975.7 (SD 224.9) face-to-face consultations, 215.7 (SD 129.1) telephone consultations and 68.0 (SD 46.9) home visits per trimester.
DISTRIBUTION OF PATIENTS’ AGES

Data from face-to-face consultations, telephone consultations and home visits were combined to determine patients’ ages (Figure 1). Univariate analysis showed significant differences in all age groups (p ≤ 0.00) between trainers and trainees, except for the 25-44 and 75+ year old patient groups.

Figure 1. Proportions (%) of contacts of trainees (n=73) and trainers (n=105*), displayed per age group

*The data of nine trainers had to be disregarded; their data were biased by large numbers of influenza vaccinations that could not be traced back to the proper age group.
ORGAN SYSTEMS (ICPC CHAPTERS)

In Figure 2, the results for all participants are combined. The DNSGP-2 data seem to be similar to our results, except for female conditions (pregnancy and female genital). The LINH 2009 figures differ on circulatory and metabolic/endocrinology, the areas in which most NPs are active.

Figure 2. Mean percentages and SD (error bars) of ICPC diagnoses codes, aggregated by chapter for trainers (n=100*) and for trainees (n=71*) and reference data of DNSGP-2** and LINH 2009**

*The data of 14 trainers and two trainees were disregarded because their mean ICPC coding percentage was lower than 50%.

**The Dutch National Survey of General Practice (DNSGP-2) and Netherlands Information Network of General Practice (LINH 2009) are based on prevalences per 1000 patients.
MOST COMMON ICPC DIAGNOSES
Acute respiratory infections headed the trainees’ top-10 list (Table 1), and when they were combined with cough, sinusitis and bronchitis, and compared with the trainers’ data, the differences for upper respiratory tract infections were even more apparent. Otitis media was ranked 7th for trainees, compared with 25th (0.71%) for trainers. Hypertension topped the trainers’ list. Diabetes mellitus and depression were in the trainers’, but not in the trainees’, top 10 (nos. 21 (0.97%) and 40 (0.85%), respectively).

Table 1. Comparison of the 10 most common ICPC diagnoses made by trainer and trainee (Years 1 and 3 combined).

<table>
<thead>
<tr>
<th>No.</th>
<th>Trainee (N=71)</th>
<th>%</th>
<th>Trainer (N=100)</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Acute respiratory infection</td>
<td>5.2</td>
<td>Hypertension</td>
<td>4.5</td>
</tr>
<tr>
<td>2</td>
<td>Cough</td>
<td>2.7</td>
<td>Acute respiratory infection</td>
<td>2.5</td>
</tr>
<tr>
<td>3</td>
<td>Hypertension</td>
<td>2.3</td>
<td>Diabetes mellitus</td>
<td>1.8</td>
</tr>
<tr>
<td>4</td>
<td>Dermatomycosis</td>
<td>1.8</td>
<td>Cough</td>
<td>1.8</td>
</tr>
<tr>
<td>5</td>
<td>Low-back pain</td>
<td>1.4</td>
<td>Low-back pain</td>
<td>1.5</td>
</tr>
<tr>
<td>6</td>
<td>Cystitis or other urinary infection</td>
<td>1.4</td>
<td>Depression</td>
<td>1.5</td>
</tr>
<tr>
<td>7</td>
<td>Otitis media</td>
<td>1.4</td>
<td>Cystitis or other urinary infection</td>
<td>1.4</td>
</tr>
<tr>
<td>8</td>
<td>Acute or chronic sinusitis</td>
<td>1.4</td>
<td>Weakness/tiredness</td>
<td>1.4</td>
</tr>
<tr>
<td>9</td>
<td>Weakness/tiredness</td>
<td>1.3</td>
<td>Dermatomycosis</td>
<td>1.3</td>
</tr>
<tr>
<td>10</td>
<td>Acute bronchitis/bronchiolitis</td>
<td>1.3</td>
<td>Asthma</td>
<td>1.3</td>
</tr>
</tbody>
</table>

PROPORTIONS OF PATIENT MIX FOR SPECIFIC DISEASES
AND DISEASE CLUSTERS
Female doctors (trainees and trainers combined) saw more female conditions \( (p<0.001) \) than their male colleagues. Female trainees saw an average percentage of 6.6% female conditions as opposed to 3.9% for male trainees. For female trainers this percentage was 8.1% versus 4.8% for male trainers. Male doctors saw more male conditions \( (p=0.001) \) than their female colleagues. Male trainees saw an average percentage of 1.3% male conditions and female trainees saw 1.1%; for male trainers this was 1.9% versus 1.3% for female trainers. For more disease clusters see Table 2.
Table 2. Comparisons of the patient mixes of trainees and trainers for specific diseases and disease clusters between the first and third training year.

<table>
<thead>
<tr>
<th>ICPC Chapter/Cluster/Disease</th>
<th>First-year trainees (N=48), mean % (SD)</th>
<th>First-year trainers (N=64), mean % (SD)</th>
<th>Third-year trainees (N=23), mean % (SD)</th>
<th>Third-year trainers (N=36), mean % (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ICPC Chapters</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General</td>
<td>5.1 (1.5)</td>
<td>5.1 (1.5)</td>
<td>5.1 (1.2)</td>
<td>5.1 (2.3)</td>
</tr>
<tr>
<td>Blood(-forming)</td>
<td>0.9 (0.5)</td>
<td>1.2 (0.5)</td>
<td>1.3 (0.6)</td>
<td>1.3 (0.9)</td>
</tr>
<tr>
<td>Digestive</td>
<td>8.9 (1.7)</td>
<td>8.5 (1.3)</td>
<td>8.9 (1.9)</td>
<td>8.3 (1.7)</td>
</tr>
<tr>
<td>Eye</td>
<td>3.3 (0.8)</td>
<td>2.4 (0.6)**</td>
<td>3.3 (1.1)</td>
<td>2.7 (0.7)</td>
</tr>
<tr>
<td>Ear</td>
<td>5.7 (1.6)</td>
<td>3.8 (1.0)**</td>
<td>4.8 (1.3)</td>
<td>3.5 (0.9)**</td>
</tr>
<tr>
<td>Circulatory</td>
<td>5.6 (2.6)</td>
<td>9.0 (3.3)**</td>
<td>6.8 (2.1)</td>
<td>9.5 (2.4)**</td>
</tr>
<tr>
<td>Musculoskeletal</td>
<td>16.4 (2.7)</td>
<td>16.4 (2.7)</td>
<td>15.3 (2.2)</td>
<td>16.2 (2.5)</td>
</tr>
<tr>
<td>Neurology</td>
<td>3.0 (0.8)</td>
<td>3.1 (0.8)</td>
<td>2.7 (0.7)</td>
<td>3.2 (0.7)*</td>
</tr>
<tr>
<td>Psychiatry</td>
<td>3.7 (1.7)</td>
<td>7.0 (2.0)**</td>
<td>5.0 (1.8)</td>
<td>7.8 (3.2)**</td>
</tr>
<tr>
<td>Respiratory</td>
<td>18.3 (3.5)</td>
<td>13.8 (2.4)**</td>
<td>16.5 (4.6)</td>
<td>12.3 (2.7)**</td>
</tr>
<tr>
<td>Skin</td>
<td>16.7 (3.1)</td>
<td>12.9 (2.0)**</td>
<td>16.7 (2.9)</td>
<td>13.8 (2.8)**</td>
</tr>
<tr>
<td>Metabolic/endocrine</td>
<td>2.1 (1.1)</td>
<td>4.1 (1.8)**</td>
<td>2.8 (1.4)</td>
<td>3.8 (1.8)</td>
</tr>
<tr>
<td>Urology</td>
<td>2.4 (1.0)</td>
<td>2.9 (1.0)</td>
<td>2.7 (1.5)</td>
<td>3.0 (1.2)</td>
</tr>
<tr>
<td>Pregnancy</td>
<td>2.0 (1.2)</td>
<td>2.3 (1.6)</td>
<td>2.0 (0.9)</td>
<td>1.7 (0.9)</td>
</tr>
<tr>
<td>Female genital</td>
<td>3.9 (2.3)</td>
<td>4.3 (1.8)</td>
<td>4.1 (1.2)</td>
<td>3.9 (1.4)</td>
</tr>
<tr>
<td>Male genital</td>
<td>1.1 (0.5)</td>
<td>1.5 (0.6)**</td>
<td>1.2 (1.0)</td>
<td>1.9 (1.1)</td>
</tr>
<tr>
<td>Social problems</td>
<td>0.8 (0.5)</td>
<td>1.7 (1.1)**</td>
<td>0.8 (0.4)</td>
<td>1.9 (1.1)**</td>
</tr>
<tr>
<td><strong>Cluster</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acute diseases a</td>
<td>1.8 (0.7)</td>
<td>2.3 (0.8)**</td>
<td>2.2 (0.9)</td>
<td>2.4 (0.8)</td>
</tr>
<tr>
<td>Chronic diseases</td>
<td>8.7 (3.2)</td>
<td>15.8 (5.3)**</td>
<td>10.8 (3.4)</td>
<td>16.2 (4.6)**</td>
</tr>
<tr>
<td>Oncological</td>
<td>0.4 (0.4)</td>
<td>1.7 (1.4)**</td>
<td>0.9 (0.9)</td>
<td>1.8 (1.4)*</td>
</tr>
<tr>
<td>Diabetes mellitus</td>
<td>0.8 (0.8)</td>
<td>1.9 (1.3)**</td>
<td>1.2 (1.3)</td>
<td>1.7 (1.4)</td>
</tr>
<tr>
<td>CVRM-primary b (not diabetes)</td>
<td>2.8 (1.7)</td>
<td>5.8 (2.4)**</td>
<td>3.7 (1.5)</td>
<td>6.2 (2.2)**</td>
</tr>
<tr>
<td>CVRM-secondary c</td>
<td>0.3 (0.2)</td>
<td>0.7 (0.5)**</td>
<td>0.4 (0.5)</td>
<td>0.8 (0.5)</td>
</tr>
<tr>
<td>COPD</td>
<td>0.6 (0.6)</td>
<td>0.9 (0.7)</td>
<td>0.7 (0.5)</td>
<td>0.7 (0.5)</td>
</tr>
<tr>
<td>Minor illnesses a</td>
<td>36.3 (5.9)</td>
<td>27.8 (4.8)**</td>
<td>35.9 (6.4)</td>
<td>28.8 (5.9)**</td>
</tr>
<tr>
<td>Symptom diagnosis</td>
<td>46.2 (6.9)</td>
<td>40.4 (6.8)**</td>
<td>43.1 (7.5)</td>
<td>39.9 (6.8)</td>
</tr>
<tr>
<td><strong>Other diseases</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypothyroiditis</td>
<td>0.1 (0.2)</td>
<td>0.4 (0.3)**</td>
<td>0.2 (0.2)</td>
<td>0.4 (0.3)</td>
</tr>
<tr>
<td>Asthma</td>
<td>1.1 (0.7)</td>
<td>1.3 (0.6)</td>
<td>1.1 (0.6)</td>
<td>1.3 (0.7)</td>
</tr>
<tr>
<td>Depression or depressed feelings</td>
<td>0.8 (0.6)</td>
<td>1.8 (0.8)**</td>
<td>1.0 (0.8)</td>
<td>1.8 (1.0)*</td>
</tr>
<tr>
<td>Irritable bowel syndrome</td>
<td>0.3 (0.3)</td>
<td>0.5 (0.4)*</td>
<td>0.4 (0.4)</td>
<td>0.5 (0.4)</td>
</tr>
<tr>
<td>Conjunctivitis</td>
<td>0.7 (0.5)</td>
<td>0.4 (0.2)**</td>
<td>0.9 (0.6)</td>
<td>0.5 (0.3)**</td>
</tr>
<tr>
<td>Sore throat d</td>
<td>1.7 (0.8)</td>
<td>1.0 (0.5)**</td>
<td>1.6 (0.7)</td>
<td>0.9 (0.4)**</td>
</tr>
</tbody>
</table>
Differences between trainers and trainees, \( p < 0.01 \).

First- and third-year trainees were tested separately against their trainers.

** \( p < 0.001 \).

a. The ICPC codes of these clusters are displayed in the supplementary appendices.

b. Cardiovascular risk management, primary prevention: problems labelled as diagnostic, screening and preventative procedures of the cardiovascular system, complicated or uncomplicated hypertension, atherosclerosis, tobacco abuse, obesity, overweight and lipid disorder.

c. Cardiovascular risk management, secondary prevention: ischaemic heart disease (with or without angina), myocardial infarction, transient ischaemic attack, stroke, intermittent claudication, aortic aneurysm. This cluster was derived from the Advice on the Recording of CVRM of the Dutch College of General Practitioners (NHG).

d. Sore throat, tonsillitis, pharyngitis, peritonsillar abscess.

COPD: chronic obstructive pulmonary disease; CVRM: cardiovascular risk management; ICPC: International Classification of Primary Care

Discussion

Seventy-three GP-training practices were investigated. Trainees saw a higher percentage of young patients, but fewer cardiovascular risk management (CVRM)-primary prevention, chronic and oncological conditions, and fewer circulatory and psychiatric diseases, social problems and neurological conditions (third year only) than their trainers. They saw more skin, ear and respiratory diseases and minor illnesses. In the first training year, they saw noticeably fewer acute diseases, and male genital and metabolic diseases (specifically diabetes) and CVRM (secondary prevention), but more eye diseases.

Similar to other studies addressing trainee patient mix,\textsuperscript{15,18,20} trainees saw more minor illnesses than their trainers. This is notable but not necessarily worrisome. Boredom and ‘saturation’ could be a concern, but are unlikely since the social background of the patients usually offers sufficient challenges for the trainees. For instance, by seeing children together with their parent(s) a trainee becomes familiar with the role of the family doctor. In addition, the most frequently encountered disease (acute respiratory infection) constituted only 5.2\% of the trainees’ patient mix (Table 1).

As in other studies,\textsuperscript{15,18,20} trainees saw fewer chronic and oncological\textsuperscript{15} conditions. This is interesting, since Dare\textsuperscript{30} found that the majority of US GPs felt inadequately trained for chronic diseases, and Card\textsuperscript{42} reported that recently graduated internists felt insufficiently prepared for chronic care. As previously reported,\textsuperscript{20,23,24,26,43} the trainees saw fewer psychosocial conditions. Given these findings and the multimorbidity of the ageing population,\textsuperscript{44} we believe that trainees should handle more severe, chronic and complex patients, and an intervention to shift the
patient mix towards the level of the trainer is worth considering.

Eccles\(^\text{18}\) Fleming\(^\text{24}\) and others\(^{14;23;26;28}\) found that male and female trainees saw fewer female conditions than their trainers. Our results show that the differences regarding female conditions were due to the doctors’ gender, as also found by Levy in a 3-week family medicine preceptorship.\(^{45}\) This might have implications for the competence building of male trainees.

Differences found in previous studies regarding circulatory,\(^{14;20;21;23;24;26;28}\) metabolic,\(^{20;23;24;26}\) skin,\(^{20;21;23;24;26}\) eye,\(^{20;26}\) ear\(^{26}\) (in most studies, the sense organs were combined with the nervous system) and respiratory diseases\(^{14;21;23;24;26}\) were confirmed in our study.

Generally, where a significant difference was found between the mean percentages of ICPC diagnosis codes encountered by the trainees and their trainers (Figure 2), the DNSGP-2 percentages were between those of the trainers and those of the trainees. Therefore, the trainees’ percentage was closer to that of the average GP than to that of the trainers. Apparently, patient supply is divided unequally, in the majority of the ICPC chapters. This uneven division was seen less often when compared to LINH 2009.

Most data were collected in the common-cold season, explaining the higher proportion of respiratory diseases. Trainees probably see most of the seasonal surplus because they usually treat patients on a short-term basis.

Pregnancy tests, contraceptive injections and PAP smears carried out by medical receptionists were disregarded in our data, whereas the DNSGP-2 included these. This might explain the differences regarding female conditions. However, the LINH online figures for 2002-2009 on female conditions were in the same range as our data.\(^{40}\)

The fact that trainers saw more patients than trainees can be partly explained by their working hours. Closer examination showed that the trainees’ workload increased over time.

Differences in patient mix may be caused by patients’ preferences for a specific doctor. Patients do not always want to consult a trainee for various reasons, for instance chronic or emotional problems were found to be associated with reduced willingness to consult registrars.\(^{46}\) Also, in another study, medical receptionists reported that patients with complex conditions were relatively more often assigned to the trainer.\(^{31}\)

Offering the trainees an adequate patient mix can be regarded as an educa-
tional obligation. Nevertheless, whether or not to assign patients with chronic or complex conditions to the trainee should be carefully weighed by trainers who can experience a conflict of interest since they are also responsible for the quality of healthcare provided. Dealing with this conflict is a professional challenge.

This study was unable to relate what the trainees saw to how much they learnt. Despite this, detailed knowledge of the trainees’ patient mix has important educational implications when the curriculum is being planned. However, disparities in patient mix between trainers and trainees may not actually be problematic since several studies have shown that supervision is of crucial importance. In case of scarcity, it might be possible to build competence despite a lower level of exposure. Trainees and trainers should therefore be aware of areas of low exposure and disparity. Trainers could increase the number of patients seen by the trainees in these areas or provide special supervised sessions for low-exposure conditions. Trainees should build expert performance in the areas they will frequently encounter. For low-exposure areas, learning aims should be to achieve minimum competence, concentrating on initial treatment and not missing important diagnoses.

Data obtained from EPR can be used to identify features of training practices, such as the attendance of many elderly patients or young children. Thus, when trainees are being allocated, their learning objectives can be matched to the practice features.

LIMITATIONS
Since the patient supply in training practices is shared between trainer and trainee, the trainer’s patient mix might vary from that normally encountered by an average GP. We therefore also compared our data with external data, which showed several apparent similarities, even though the data are not fully comparable.

The assumption that trainers or other GPs have an adequate patient mix is debatable. Alternatively, the trainees’ patient mix can be compared to standards used in GP specialty training. These are consensus standards and largely based on national morbidity figures and therefore rest on the same assumption. From an educational perspective, these standards lack a scientific basis. More specific evidence is needed concerning the relationship between patient mix and learning so that patient mix can be optimised.

Trainers with a coding percentage over 90% saw more psychiatry and social problems than trainers in the 50%-90% range (not reported), indicating that psy-
chosocial codes are under-reported by some trainers. This was not found for trainees.

Although ICPC coding is useful for describing patient mix, its validity is potentially diminished by its dependence on the diagnostic competence of doctors and the accurate attribution of codes (eg, definite instead of symptom diagnosis).

All participants were affiliated to one training institute. However, we do not think that this influenced the results, as other, different studies have shown similar results.

Minor illnesses and acute diseases were interesting as regards the description of learning experiences. Previous authors reporting on these clusters were implicit about their contents.\textsuperscript{15,18,20} We therefore created these clusters ourselves, but only with medium methodological rigour. Future studies should construct valid, sound thematic clusters which have educational meaning.

CONCLUSION

We studied the patient mix of GP-training practices using EPR on a large scale and over a long period and confirmed that the disparities between the patient mixes of trainers and trainees are considerable but stable over time and location. The difference in contact frequencies needs attention, as does exposure to chronic, psychosocial and circulatory diseases, and, probably, severe and complex conditions. Trainees should be given the opportunity to handle these conditions more frequently. Future studies should focus on interventions to tailor the patient mix, by, for instance, focusing on low-exposure areas. Individual trainee’s patient mix should be adjusted to their personal development plan. EPR systems are excellent for permanently monitoring this patient mix, and thereby the trainee’s learning experiences.
REFERENCES


41. van der Linden M, Westert GP, Schellevis FG, et al. Tweede Nationale Studie: naar ziekten en verrichtingen in de huisartspraktijk. (Second National Study into Diseases and Actions in General Practice). Utrecht (Bilthoven): Dutch Institute for Health Services Research (NIVEL) and National


### Appendix 1.

#### ACUTE CONDITIONS

<table>
<thead>
<tr>
<th>ICD-10 Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A02</td>
<td>Chills</td>
</tr>
<tr>
<td>A07</td>
<td>Coma</td>
</tr>
<tr>
<td>A10</td>
<td>Bleeding/haemorrhage nos*</td>
</tr>
<tr>
<td>A12.1</td>
<td>Anaphylactic shock</td>
</tr>
<tr>
<td>A12.2</td>
<td>Angioedema</td>
</tr>
<tr>
<td>A81</td>
<td>Multiple trauma/injuries</td>
</tr>
<tr>
<td>A84</td>
<td>Poisoning by medical agent</td>
</tr>
<tr>
<td>A86</td>
<td>Toxic effect non-medicinal substance</td>
</tr>
<tr>
<td>A96.2</td>
<td>Unnatural death</td>
</tr>
<tr>
<td>B76</td>
<td>Ruptured spleen traumatic</td>
</tr>
<tr>
<td>D14</td>
<td>Haematemesis/vomiting blood</td>
</tr>
<tr>
<td>D15</td>
<td>Melaema</td>
</tr>
<tr>
<td>D88</td>
<td>Appendicitis</td>
</tr>
<tr>
<td>D92</td>
<td>Diverticular disease</td>
</tr>
<tr>
<td>D98.1</td>
<td>Cholecystitis</td>
</tr>
<tr>
<td>D98.2</td>
<td>Cholangitis</td>
</tr>
<tr>
<td>D99.1</td>
<td>Ileus</td>
</tr>
<tr>
<td>D99.2</td>
<td>Intussusception</td>
</tr>
<tr>
<td>D99.3</td>
<td>Peritonitis</td>
</tr>
<tr>
<td>D99.4</td>
<td>Pancreatitis</td>
</tr>
<tr>
<td>F76</td>
<td>Foreign body in eye</td>
</tr>
<tr>
<td>F79</td>
<td>Injury eye other</td>
</tr>
<tr>
<td>F82</td>
<td>Detached retina</td>
</tr>
<tr>
<td>F85.1</td>
<td>Herpetic keratitis</td>
</tr>
<tr>
<td>F83</td>
<td>Glaucoma</td>
</tr>
<tr>
<td>F83.2</td>
<td>Glaucoma, acute angle closure</td>
</tr>
<tr>
<td>H71.2</td>
<td>Mastoiditis</td>
</tr>
<tr>
<td>K01</td>
<td>Heart pain</td>
</tr>
<tr>
<td>K02</td>
<td>Pressure/tightness of heart</td>
</tr>
<tr>
<td>K03</td>
<td>Cardiovascular pain nos*</td>
</tr>
<tr>
<td>K71.2</td>
<td>Acute rheumatic fever with heart disease</td>
</tr>
<tr>
<td>K74.1</td>
<td>Unstable angina</td>
</tr>
<tr>
<td>K75</td>
<td>Acute myocardial infarction</td>
</tr>
<tr>
<td>K77.1</td>
<td>Acute decompensated heart failure</td>
</tr>
<tr>
<td>K78.2</td>
<td>Ventricular tachycardia</td>
</tr>
<tr>
<td>K80.3</td>
<td>Sick sinus syndrome</td>
</tr>
<tr>
<td>K89</td>
<td>Transient cerebral ischaemia</td>
</tr>
<tr>
<td>K90</td>
<td>Stroke/cerebrovascular accident</td>
</tr>
<tr>
<td>K90.1</td>
<td>Subarachnoid hemorrhage</td>
</tr>
<tr>
<td>K90.2</td>
<td>Cerebral hemorrhage</td>
</tr>
<tr>
<td>K90.3</td>
<td>Cerebral infarction</td>
</tr>
<tr>
<td>K93</td>
<td>Pulmonary embolism</td>
</tr>
<tr>
<td>K94.1</td>
<td>Deep venous thrombosis</td>
</tr>
<tr>
<td>K98.1</td>
<td>Aortic aneurysm</td>
</tr>
<tr>
<td>K99.5</td>
<td>Giant-cell arteritis</td>
</tr>
<tr>
<td>L70.2</td>
<td>Septic arthritis</td>
</tr>
<tr>
<td>L72</td>
<td>Fracture: radius/ulna</td>
</tr>
<tr>
<td>L73</td>
<td>Fracture: tibia/fibula</td>
</tr>
<tr>
<td>L74</td>
<td>Fracture: hand/foot bone</td>
</tr>
<tr>
<td>L75</td>
<td>Fracture: femur</td>
</tr>
<tr>
<td>L75.1</td>
<td>Hip fracture</td>
</tr>
<tr>
<td>L75.2</td>
<td>Skull fracture</td>
</tr>
<tr>
<td>L76.4</td>
<td>Humerus fracture</td>
</tr>
<tr>
<td>L76.6</td>
<td>Spinal fracture</td>
</tr>
<tr>
<td>L76.7</td>
<td>Pelvic fracture</td>
</tr>
<tr>
<td>L76.8</td>
<td>Patella fracture</td>
</tr>
<tr>
<td>L80.2</td>
<td>Jaw dislocation</td>
</tr>
<tr>
<td>L80.3</td>
<td>Finger dislocation</td>
</tr>
<tr>
<td>L80.4</td>
<td>Acromioclavicular separation</td>
</tr>
<tr>
<td>N07.1</td>
<td>Fracture: radius/ulna</td>
</tr>
<tr>
<td>N18</td>
<td>Paralysis/weakness</td>
</tr>
<tr>
<td>N71.1</td>
<td>Bacterial meningitis</td>
</tr>
<tr>
<td>N71.2</td>
<td>Viral meningitis</td>
</tr>
<tr>
<td>N71.3</td>
<td>Encephalitis</td>
</tr>
<tr>
<td>N71.4</td>
<td>Myelitis</td>
</tr>
<tr>
<td>N72</td>
<td>Tetanus</td>
</tr>
<tr>
<td>N80.1</td>
<td>Posttraumatic epidural hematoma</td>
</tr>
<tr>
<td>N80.2</td>
<td>Posttraumatic subdural hematoma</td>
</tr>
<tr>
<td>N80.3</td>
<td>Posttraumatic intracerebral hematoma</td>
</tr>
<tr>
<td>N80.4</td>
<td>Brain contusion</td>
</tr>
<tr>
<td>N88</td>
<td>Epilepsy</td>
</tr>
<tr>
<td>N94.1</td>
<td>Guillain-barré syndrome</td>
</tr>
<tr>
<td>P15.2</td>
<td>Delirium tremens</td>
</tr>
<tr>
<td>P16</td>
<td>Acute alcohol abuse</td>
</tr>
<tr>
<td>P71</td>
<td>Organic psychosis other</td>
</tr>
<tr>
<td>P77</td>
<td>Suicide/suicide attempt</td>
</tr>
<tr>
<td>P77.1</td>
<td>Attempted suicide</td>
</tr>
<tr>
<td>P98</td>
<td>Psychosis nos*/other</td>
</tr>
<tr>
<td>R77.2</td>
<td>Suicide</td>
</tr>
<tr>
<td>R81.1</td>
<td>Legionella pneumonia</td>
</tr>
<tr>
<td>R83.1</td>
<td>Diptheria</td>
</tr>
<tr>
<td>R87</td>
<td>Foreign body nose/larynx/bronch</td>
</tr>
<tr>
<td>R99.4</td>
<td>Pneumothorax</td>
</tr>
<tr>
<td>R99.5</td>
<td>Aspiration pneumonia</td>
</tr>
<tr>
<td>R99.6</td>
<td>Pneumococciosis</td>
</tr>
<tr>
<td>S14</td>
<td>Burn/scald</td>
</tr>
<tr>
<td>T11</td>
<td>Dehydration</td>
</tr>
<tr>
<td>T87</td>
<td>Hypoglycaemia</td>
</tr>
<tr>
<td>U4.1</td>
<td>Renal colic</td>
</tr>
<tr>
<td>U70</td>
<td>Pyelonephritis/pyelitis</td>
</tr>
<tr>
<td>U80.1</td>
<td>Renal contusion</td>
</tr>
<tr>
<td>U80.2</td>
<td>Foreign body urinary tract</td>
</tr>
<tr>
<td>U95</td>
<td>Urinary calculus</td>
</tr>
<tr>
<td>W03.2</td>
<td>Bleeding during second/third trimester</td>
</tr>
<tr>
<td>W29.1</td>
<td>Decreased fetal movement</td>
</tr>
<tr>
<td>W70.1</td>
<td>Postpartum endometritis</td>
</tr>
<tr>
<td>W77.3</td>
<td>DVT in pregnancy</td>
</tr>
<tr>
<td>W80</td>
<td>Ectopic pregnancy</td>
</tr>
<tr>
<td>W81.2</td>
<td>Toxaemia of pregnancy</td>
</tr>
<tr>
<td>W81.3</td>
<td>HELLP syndrome</td>
</tr>
<tr>
<td>W82.1</td>
<td>Caesarean section</td>
</tr>
<tr>
<td>W82.2</td>
<td>Vacuum extraction</td>
</tr>
<tr>
<td>W82.3</td>
<td>Forcipal extraction</td>
</tr>
<tr>
<td>W93</td>
<td>Complicate labour/delivery stillbirth</td>
</tr>
<tr>
<td>W99.3</td>
<td>Postnatal thrombosis</td>
</tr>
<tr>
<td>Y99.1</td>
<td>Testicular torsion</td>
</tr>
</tbody>
</table>

*Not Otherwise Specified*
## Appendix 2.
### Minor Illnesses

<table>
<thead>
<tr>
<th>ICD Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A09</td>
<td>Sweating problem</td>
</tr>
<tr>
<td>A09.2</td>
<td>Focal excessive sweating</td>
</tr>
<tr>
<td>A72</td>
<td>Chickenpox</td>
</tr>
<tr>
<td>A75</td>
<td>Infectious mononucleosis</td>
</tr>
<tr>
<td>A76</td>
<td>Viral exanthem other</td>
</tr>
<tr>
<td>A76.1</td>
<td>Exanthema subitum/sixth disease</td>
</tr>
<tr>
<td>A76.2</td>
<td>Erythema infectiosum / fifth disease</td>
</tr>
<tr>
<td>A76.3</td>
<td>Hand, foot and mouth disease</td>
</tr>
<tr>
<td>A85.2</td>
<td>Spotting with oral contraception use</td>
</tr>
<tr>
<td>A88.1</td>
<td>Chilblains</td>
</tr>
<tr>
<td>A88.3</td>
<td>Motion sickness</td>
</tr>
<tr>
<td>B70</td>
<td>Lymphadenitis acute</td>
</tr>
<tr>
<td>B71</td>
<td>Lymphadenitis non-specific</td>
</tr>
<tr>
<td>D05</td>
<td>Perianal itching</td>
</tr>
<tr>
<td>D08</td>
<td>Flatulence/gas/belching</td>
</tr>
<tr>
<td>D19</td>
<td>Teeth/gum symptom/complaint</td>
</tr>
<tr>
<td>D20</td>
<td>Mouth/tongue/lip symptom/complaint</td>
</tr>
<tr>
<td>D22</td>
<td>Helminths/oxyuriasis / other parasites</td>
</tr>
<tr>
<td>D22.1</td>
<td>Enterobiasis</td>
</tr>
<tr>
<td>D22.2</td>
<td>Ascariasis</td>
</tr>
<tr>
<td>D22.3</td>
<td>Taeniasis</td>
</tr>
<tr>
<td>D70.3</td>
<td>Giardiasis</td>
</tr>
<tr>
<td>D73</td>
<td>Gastroenteritis presumed infection</td>
</tr>
<tr>
<td>D81.2</td>
<td>Tongue-tie</td>
</tr>
<tr>
<td>D81.2</td>
<td>Gingivitis</td>
</tr>
<tr>
<td>D82.2</td>
<td>Temporomandibular joint disorder</td>
</tr>
<tr>
<td>D83.1</td>
<td>Sialoadenitis/sialolithiasis</td>
</tr>
<tr>
<td>D83.2</td>
<td>Aphthous ulcer</td>
</tr>
<tr>
<td>D83.3</td>
<td>Perleche/ angular cheilitis</td>
</tr>
<tr>
<td>D84.2</td>
<td>Gastroesophageal reflux disease without oesophagitis</td>
</tr>
<tr>
<td>D872</td>
<td>Dyspepsia/indigestion</td>
</tr>
<tr>
<td>D93</td>
<td>Irritable bowel syndrome</td>
</tr>
<tr>
<td>D95</td>
<td>Anal fissure/perianal abscess</td>
</tr>
<tr>
<td>D95.1</td>
<td>Anal fissure</td>
</tr>
<tr>
<td>F03</td>
<td>Eye discharge</td>
</tr>
<tr>
<td>F03.1</td>
<td>Watery eyes</td>
</tr>
<tr>
<td>F04</td>
<td>Visual floaters/spots</td>
</tr>
<tr>
<td>F16</td>
<td>Eyelid symptom/complaint</td>
</tr>
<tr>
<td>F17</td>
<td>Glasses symptom/complaint</td>
</tr>
<tr>
<td>F18</td>
<td>Contact lens symptom/complaint</td>
</tr>
<tr>
<td>F70.1</td>
<td>Conjunctivitis infectious</td>
</tr>
<tr>
<td>F70.2</td>
<td>Viral conjunctivitis</td>
</tr>
<tr>
<td>F71</td>
<td>Conjunctivitis allergic</td>
</tr>
<tr>
<td>F71.1</td>
<td>Allergic conjunctivitis</td>
</tr>
<tr>
<td>F72.1</td>
<td>Blepharitis</td>
</tr>
<tr>
<td>F72.2</td>
<td>Stye</td>
</tr>
<tr>
<td>F72.3</td>
<td>Chalazion</td>
</tr>
<tr>
<td>F73.1</td>
<td>Dacryocystitis</td>
</tr>
<tr>
<td>F75.1</td>
<td>Subconjunctival hemorrhage</td>
</tr>
<tr>
<td>F75.2</td>
<td>Black eye</td>
</tr>
<tr>
<td>F76</td>
<td>Foreign body in eye</td>
</tr>
<tr>
<td>F79.1</td>
<td>Photokeratitis</td>
</tr>
<tr>
<td>F79.2</td>
<td>Corneal erosion</td>
</tr>
<tr>
<td>F80</td>
<td>Blocked lacrimal duct of infant</td>
</tr>
<tr>
<td>F91</td>
<td>Refractive error</td>
</tr>
<tr>
<td>F91.1</td>
<td>Presbyopia</td>
</tr>
<tr>
<td>F91.2</td>
<td>Hyperopia</td>
</tr>
<tr>
<td>F91.3</td>
<td>Myopia</td>
</tr>
<tr>
<td>F91.4</td>
<td>Astigmatism</td>
</tr>
<tr>
<td>F98.1</td>
<td>Xanthelasma palpebrarum</td>
</tr>
<tr>
<td>F98.2</td>
<td>Blepharochalasis</td>
</tr>
<tr>
<td>F99.3</td>
<td>Entropion/ectropion</td>
</tr>
<tr>
<td>F99.4</td>
<td>Pterygium</td>
</tr>
<tr>
<td>F99.6</td>
<td>Insufficient tear film</td>
</tr>
<tr>
<td>H01</td>
<td>Ear pain/earache</td>
</tr>
<tr>
<td>H02</td>
<td>Hearing complaint</td>
</tr>
<tr>
<td>H03</td>
<td>Tinnitus, ringing/buzzing ear</td>
</tr>
<tr>
<td>H04</td>
<td>Ear discharge</td>
</tr>
<tr>
<td>H05</td>
<td>Plugged feeling ear</td>
</tr>
<tr>
<td>H06</td>
<td>Concern with appearance of ears</td>
</tr>
<tr>
<td>H06.1</td>
<td>Protruding ear</td>
</tr>
<tr>
<td>H07</td>
<td>Otitis externa</td>
</tr>
<tr>
<td>H07.1</td>
<td>Acute otitis media/myringitis</td>
</tr>
<tr>
<td>H07.2</td>
<td>Serous otitis media</td>
</tr>
<tr>
<td>H07.3</td>
<td>Eustachian salpingitis</td>
</tr>
<tr>
<td>H07.6</td>
<td>Foreign body in ear</td>
</tr>
<tr>
<td>H07.7</td>
<td>Perforation ear drum</td>
</tr>
<tr>
<td>H07.8</td>
<td>Superficial injury of ear</td>
</tr>
<tr>
<td>H08</td>
<td>Excessive ear wax</td>
</tr>
<tr>
<td>H08.2</td>
<td>Vertiginous syndrome</td>
</tr>
<tr>
<td>H08.2</td>
<td>Labyrinthitis</td>
</tr>
<tr>
<td>H08.3</td>
<td>Paroxysmal positional vertigo</td>
</tr>
<tr>
<td>K06</td>
<td>Prominent veins</td>
</tr>
<tr>
<td>K08.1</td>
<td>Cardiac arrhythmia nos*</td>
</tr>
<tr>
<td>K08.2</td>
<td>Supraventricular extrasystole</td>
</tr>
<tr>
<td>K08.8</td>
<td>Postural hypotension</td>
</tr>
<tr>
<td>K09.2</td>
<td>Raynaud's phenomenon</td>
</tr>
<tr>
<td>K09.5</td>
<td>Varicose veins of leg</td>
</tr>
<tr>
<td>K09.6</td>
<td>Haemorrhoids</td>
</tr>
<tr>
<td>K09.9.2</td>
<td>Varicocle</td>
</tr>
<tr>
<td>K99.4</td>
<td>Chronic venous insufficiency</td>
</tr>
<tr>
<td>L06</td>
<td>Axilla symptoms/complaints</td>
</tr>
<tr>
<td>L08</td>
<td>Muscle pain</td>
</tr>
<tr>
<td>L19</td>
<td>Muscle symptom/complaint nos*</td>
</tr>
<tr>
<td>L77</td>
<td>Sprain/strain of ankle</td>
</tr>
<tr>
<td>L78</td>
<td>Sprain/strain of knee</td>
</tr>
<tr>
<td>L79</td>
<td>Sprain/strain of joint nos*</td>
</tr>
<tr>
<td>L80.5</td>
<td>Radial head subluxation/nursemaid's elbow</td>
</tr>
<tr>
<td>L81.1</td>
<td>Calf muscle strain/tennis leg</td>
</tr>
<tr>
<td>L81.2</td>
<td>Rib contusion</td>
</tr>
<tr>
<td>L84.1</td>
<td>Osteoarthritis/spondylosis</td>
</tr>
<tr>
<td>L87</td>
<td>Bursitis/tennisitis/synovitis nos*</td>
</tr>
<tr>
<td>L82</td>
<td>Shoulder syndrome</td>
</tr>
<tr>
<td>L93</td>
<td>Tennis elbow</td>
</tr>
<tr>
<td>L94.2</td>
<td>Osgood–schlatter disease</td>
</tr>
<tr>
<td>L98.1</td>
<td>Mallet finger</td>
</tr>
<tr>
<td>L98.2</td>
<td>Flat feet</td>
</tr>
</tbody>
</table>
chapter 5

L98.3 Bunion
L98.4 Hammer toe
L98.1 Bursitis
L98.2 Tendinitis
L98.3 Dupuytren's contracture
L98.4 Trigger finger
L98.5 Medial epicondylitis
L98.6 Tietze syndrome
L98.7 Patellofemoral pain syndrome
L98.8 Heel spur/plantar fasciitis
L98.9 Hypermobility
N02 Tension headache
N04 Restless legs
N05 Tingling fingers/feet/toes
N17.2 Lightheadedness
N89 Migraine
N93 Carpal tunnel syndrome
N94.3 Thoracic outlet syndrome
N94.4 Meralgia paraesthetica
N94.6 Morton's metatarsalgia
P04 Feeling/behaving irritable/angry
P05 Senility, feeling/behaving old
P07 Sexual desire reduced
P10 Stammering/stuttering/tic
P10.1 Stammering/stuttering
P10.2 Tic/stereotyped movements
P12 Bedwetting/enuresis
R04.1 Snoring
R06 Nose bleed/epistaxis
R07 Sneezing/nasal congestion
R09 Sinus symptom/complaint
R21 Throat symptom/complaint
R21.1 Sore throat
R22 Tonsils symptoms/complaint
R72 Strep throat
R72.1 Strep throat
R72.2 Scarlet fever
R73 Boil/abscess nose
R74 Upper respiratory infection acute
R74.1 Common cold
R74.2 Acute pharyngitis
R75 Sinusitis acute/chronic
R75.1 Acute sinusitis
R75.2 Chronic sinusitis
R76 Tonsillitis acute/peritonsillar abscess
R76.1 Tonsillitis acute
R77 Laryngitis/tracheitis acute
R77.1 Group/laryngotracheobronchitis
R78 Acute bronchitis/bronchiolitis
R80 Influenza
R97 Hay fever/allergic rhinitis
R98 Hyperventilation syndrome
R99.1 Nasal septum deviation
R99.2 Nasal polyps
R99.3 Vocal cord nodules
S02 Pruritus
S03 Warts
S09 Infected finger/toe
S09.1 Paronychia
S10 Boil/carbuncle/vesicle
S10.1 Boil/carbuncle
S10.2 Recurrent cutaneous abscesses
S10.3 Cellulitis
S11 Folliculitis
S12 Insect bite/sting
S13 Animal/human bite
S15 Foreign body in skin
S16 Bruise/contusion
S16.1 Subungual hematoma
S17 Abrasion/scratch/blister
S18 Laceration/cut
S19.1 Nail avulsion
S20 Corn/callosity
S21.1 Xeroderma
S22 Nail symptom/complaint
S22.2 Lichenification/induration
S23 Hair loss/baldness
S23.1 Alopecia areata
S23.2 Androgenic alopecia
S24.1 Hirsutism
S24.2 Change of hair color
S70 Herpes zoster
S71 Herpes simplex
S71.1 Herpes labialis
S72 Scabies/other acaraisis
S73.1 Pediculosis capitis
S73.2 Pediculosis pubis
S74.1 Athlete's foot
S74.2 Onychomycosis
S74.3 Tinea versicolor
S75.1 Oral candidiasis/thrush
S75.2 Cuticle candidiasis
S75.3 Intertigo/diaper candidiasis
S76.1 Erysipelas
S76.2 Erythrasma
S78 Lipoma
S79.1 Benign fibrous histiocytomas
S81 Haemangioma/lymphangioma
S82 Naevus/mole
S83.1 Naevus flammeus
S83.2 Capillary hemangioma
S84 Impetigo
S85 Pilonidal cyst/fistula
S86.1 Dandruff
S86.2 Cradle cap
S87 Dermatitis/atopic eczema
S88.1 Contact dermatitis
S88.2 Dyshidrosis
S88.3 Dermatitis after local drug application
S88.4 Photodermatosi
S89 Diaper rash
S90 Pityriasis rosea
S92 Sweat gland disease
S92.1 Anhidrosis/dyshidrosis
S92.2 Hidradenitis
S93.1 Sebaceous cyst
S93.2 Epithelial cyst
S94.1 Ingrown toenail
S94.2 Onychogryphosis
S95 Molluscum contagiosum
S96 Acne
S96.1 Acne vulgaris
S96.2 Acne conglobata
S98 Urticaria
S99.1 Keratoacanthoma
S99.2 Seborrhoeic keratosis
S99.3 Rosacea
S99.4 Vitiligo
S99.6 Lichen planus
S99.7 Stretch marks
T04 Feeding problem of infant/child
T81 Goitre
T92 Gout
T92.3 Lactose intolerance
U04.1 Stress incontinence
U71.1 Cystitis
U72 Urethritis
W29.2 Dissatisfied with appearance during pregnancy
W77.1 Varicose veins during pregnancy
W77.2 Haemorrhoids during pregnancy
W95.1 Breastfeeding disorders
W95.2 Cracked nipples
W99.4 Puerperal haemorrhoids
X07.2 Irregular menstruation
X08.1 Mittelschmerz
X09 Premenstrual symptom/complaint
XII.2 Atrophic vaginitis
X29.1 Dissatisfied with appearance of breasts
X72 Genital candidiasis female
X73 Genital trichomoniasis female
X78.1 Uterine fibroids
X79 Benign neoplasm breast female
X84.1 Chlamydia vaginitis
X84.2 Bacterial vaginosis
X85.2 Cervical ectropion
X85.3 Cervical polyp
X87.1 Cystocele/rectocele
X87.2 Uterine prolaps
X89 Premenstrual tension syndrome
X91 Condylomata acuminata female
X99.1 Bartholin’s cyst
Y03 Urethral discharge
Y16 Breast symptom/complaint male
Y75.1 Candida balanitis
Y76 Condylomata acuminata male
Y79 Benign/unspec. Neoplasm gen. (m)
Y81 Phimosis/redundant prepuce
Y84.1 Retractile testes
Y85 Benign prostatic hypertrophy
Y86 Hydrocoele

* Not Otherwise Specified
Who determines the patient mix of GP trainees? The role of the receptionist

Jip de Jong, Mechteld R.M. Visser and Margreet Wieringa-de Waard

Abstract

**Background** During their specialty training, Dutch GP trainees work at a GP under the supervision of a GP trainer. Research suggests that the patient mix of GP trainees differs from that of their trainers. Receptionists assign patients to either the trainee or the trainer, thereby influencing the patient mix of the trainees. The decision to which doctor to assign is complex and depends on the latitude the receptionists have. Their considerations when assigning patients are unknown.

**Objective** To study receptionists’ assigning behaviour.

**Methods** This was a questionnaire survey. To design the questionnaire, topics about assigning behaviour were identified in a focus group. The resulting questionnaire was sent to 478 GP training practices in the Netherlands.

**Results** Response rate was 68%. Of the receptionists, 95% asked for the reason for the consultation at least ‘sometimes’. Most (86.3%) of the receptionists considered the patient mix of trainees and trainers to be similar. Almost all receptionists (97%) reported ‘often’ or ‘always’ assigning ‘every possible problem’ to the trainee and a similar picture arose regarding specific subpopulations. However, the receptionists reported that they assigned complex and new patients to the trainers more often than to trainees.

**Conclusion** With some exceptions, receptionists try to assign trainees a varied patient mix
Introduction

The health problems GP trainees encounter are presumed to be sufficient to gain the competencies they are expected to master by the end of their training, but little is known about the actual health problems (‘patient mix’) GP trainees are confronted with. Earlier research regarding differences between the patient mix of GP trainees and trainees suggests that the patient mix for trainees consists primarily of minor ailments and does not include sufficient chronic diseases or severe conditions.\textsuperscript{1–4} The extent to which this finding negatively influences trainees’ development is unknown. Research has shown a relationship between patient mix and learning if exposure to patients is combined with supervision or if students show a deep learning style.\textsuperscript{5–13}

Disparity in patient mix is not a problem exclusive to GP training. It was also found in undergraduate medical curricula, clerkships and other specialty training programmes in several European countries and North America.\textsuperscript{1,7,14–19} Since GP trainers facilitate training, they are expected to provide trainees with an adequate patient mix, i.e. without educational gaps and sufficiently covering the working field.\textsuperscript{20} However, it is doubtful whether the trainee is indeed provided with such a patient mix and unknown to what extent it can be influenced.

We started this study because we wanted to know whether the differences in patient mix distribution between the doctors holding surgery found in earlier studies could be explained by the assigning behaviour of GP receptionists. In the study, we discriminate three factors that influence the patient mix of trainees practice setting, surgery planning and receptionists’ assigning behaviour.

PRACTICE SETTING

Practices vary in many ways, from single-handed to group practices and health centres, the gender of the doctors, geographic location, the proximity to an emergency unit and the gender and age distribution of a population.

SURGERY PLANNING

Surgery planning refers to the procedures and agreements concerning the planning of surgery hours. In the Dutch situation, patients are registered with a permanent GP and trainees only encounter patients registered with their trainer. Patients usually make appointments by telephone or at the desk and receptionists may ask after
the reason for the consultation and any preference for consulting either the trainer or the trainee. These kinds of questions give receptionists the ability to ‘decide’ to which doctor the patient is assigned. Obviously, receptionists also consider the urgency of the problem in combination with the doctors’ schedules, existing arrangements regarding assigning, the availability of walk-in hours and unbooked surgery for emergencies. The presence of Nurse Practitioners may also influence the patient mix\textsuperscript{21–26} because they generally hold surgery for long-term medical conditions, thereby diminishing the caseload for GP trainees.

**ASSIGNING BEHAVIOUR**

Little is known about the receptionists’ considerations when assigning patients to trainers or trainees. Receptionists may consider the amount of experience a trainee has in relation to the complexity of a problem, a trainee’s specific wishes related to current learning ends, a trainee’s gender and they may have personal ideas about what a trainee should see. The decision to assign a patient to the trainee or the trainer can, at least in some cases, be complex. Since receptionists seem to be in a position to determine at least part of the patient mix, it is important to know whether their assigning behaviour promotes or inhibits an adequate patient mix for trainees.

This study was conducted to gain a better understanding of the considerations of receptionists with respect to their patient assigning behaviour. The research question was which assigning behaviour do receptionists report?

**Methods**

**SETTING**

This study was conducted in 2007 in practices affiliated to the GP specialty training of the Academic Medical Center (AMC), University of Amsterdam. This training institute facilitates a 3-year training programme in which first- and third-year trainees are stationed in GP training practices. They work under close supervision of a trainer. GP training practices are located primarily in urban and suburban areas and a small number in rural areas. In their second year, trainees do clinical rotations.

**QUESTIONNAIRE DEVELOPMENT**

The paucity of knowledge about assigning behaviour prompted us to conduct a fo-
Focus group study of receptionists to elicit possible considerations when assigning. The results served as input for a questionnaire.

**QUESTIONNAIRE**

A self-report questionnaire was developed based on the results of the focus group (see Appendix 1) and included questions about practice setting, surgery planning and assigning behaviour. Assigning behaviour was assessed in general (11 questions) and for specific subpopulations (15 questions) using five-point Likert scales (see Text box 1, for examples). The choice of subpopulations was based on the results of the focus group, too.

Text box 1. Example questions from the questionnaire

**Organization of the surgery hours**
- How often do you ask patients their reason for the consultation?
- How often do you explicitly ask patients their preference for a specific doctor?

**Assigning patients, general issues**
- I intentionally assign a varied patient mix to the surgery hours of the trainee.
- I assign patients to the surgery hours of the trainee who have consulted the trainee before.
- I assign patients with relatively easy problems to the surgery hours of the trainee.
- The decision whether to assign to the trainer or the trainee depends on the reason for the consultation.

1 = never / 2 = seldom / 3 = sometimes / 4 = often / 5 = (nearly) always.

**Assigning patient subpopulations**
- I assign these patients:
  - Emergency cases
  - Patients with severe psychological problems
  - Children
  - New patients
  - Patients with minor ailments
  - Patients with severe psychological problems
  - Difficult cases

1 = always to the trainer / 2 = preferably to the trainer / 3 = no preference / 4 = preferably to the trainee / 5 = always to the trainee
As a pilot, three receptionists filled in the questionnaire and explained their comments, after which the questionnaire was edited and finalized. The entire questionnaire can be obtained from the first author.

**PARTICIPANTS AND PROCEDURE**
The Netherlands has ~1500 GP training practices. Each training institute hand-picked every third affiliated training practice from an alphabetically ordered list and sent them the questionnaire (n = 478). We enclosed a non-responder form asking for reasons for non-response. Since the researchers were blind to the addresses of the practices, all practices received a reminder after 3 weeks, regardless of having returned the questionnaire or not. Ethical approval was not required by Dutch law, although the questionnaires returned were handled with strict confidentiality.

**DATA ANALYSIS**
Double data entry was employed and descriptive analyses were performed. To uncover patterns in receptionists’ assigning behaviour with respect to different sub-populations, we performed a principal component analysis (PCA) using Varimax rotation with Kaiser normalization. We interpreted the resulting components with an eigenvalue >1.0. Statistical analyses were performed using SPSS 14.0.

**Results**

**FOCUS GROUP**
The focus group consisted of 10 receptionists who were recruited from the participants of a training course at our institute. Their professional experience varied from 5 to 27 years. Delegates from singlehanded, duo and group practices, as well as health centres were present. The most important results of the focus group are summarized in the Appendix 1.
RESULTS FROM THE QUESTIONNAIRES

Response. Of the 478 questionnaires sent out, 326 were returned (68%). Of the 152 non-responders, only 15 returned non-responder forms.

Participants and trainees. At the time of filling in the questionnaire, respondents had worked as receptionists for an average 10.3 (SD = 7.4) years. The proportion of first- and third-year trainees they worked with was 47.5% and 52.5%, respectively. The trainees worked an average of 7.2 (SD = 3.9) months at the respondents’ practices. Two-thirds (66.9%) of the GP trainees were female, equally divided over the first and third training years. This is in accordance with national figures: in 2007, 70% of all GP trainees in the Netherlands were female.27

Practice setting and surgery planning. These results are displayed in Table 1. Of the receptionists, 67.5% ‘always’ or ‘often’ asked patients for the reason for the consultation. About a quarter (27%) of the receptionists asked this ‘sometimes’ and a minority (5%) ‘seldom’ or ‘never’ (data not shown). As far as the preference for a specific doctor is concerned, 41.3% often or always asked whether patients had a preference, 25.9% asked this sometimes and 32.7% never or seldom (data not shown).

Unbooked surgery for urgent cases was available in 67.9% of the practices. In 87.3%, this applied to the surgery hours of both the trainee and the trainer. Walk-in hours were held in 11% of the practices. During these hours, patients were usually seen by the doctor in order of arrival and the receptionist could exert no influence on the assignment.

Arrangements between the GP trainer and receptionist concerning the assignment of face-to-face consultations to the GP trainee were made in 29.4% of the practices. These arrangements were diverse and no evident pattern could be distinguished. For example, in some practices, the receptionists assigned ‘intra-uterine device insertions’ to the GP trainee or ‘easy cases only’.

Half (45.5%) of the receptionists reported arrangements for assigning home visits to GP trainees. These arrangements were even more diverse, ranging from ‘no care of terminal patients’ to ‘all emergencies’.

In 63.9% of the practices, meetings between the GP’s and the receptionist about surgery planning were held on a daily (37.1%), weekly (17.6%) or monthly basis (9.3%).
Table 1: Practice setting and surgery planning

<table>
<thead>
<tr>
<th>Factor</th>
<th>% (yes)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Practice setting</strong></td>
<td></td>
</tr>
<tr>
<td>Single-handed practice</td>
<td>29.1</td>
</tr>
<tr>
<td>Male GP trainer</td>
<td>78.1</td>
</tr>
<tr>
<td><strong>Surgery planning</strong></td>
<td></td>
</tr>
<tr>
<td>Receptionist asked patients their reason for the consultation*</td>
<td>67.5</td>
</tr>
<tr>
<td>Receptionists asked patients for the preference for a specific doctor*</td>
<td>41.3</td>
</tr>
<tr>
<td>Unbooked surgery available in the schedule**</td>
<td>67.9</td>
</tr>
<tr>
<td>Arrangements of GP trainer and receptionist about consultations for trainee**</td>
<td>29.4</td>
</tr>
<tr>
<td>Arrangements of GP trainer and receptionist about home visits for trainee**</td>
<td>45.4</td>
</tr>
<tr>
<td>Walk-in hours**</td>
<td>11.0</td>
</tr>
<tr>
<td>Meetings about assigning (GP and receptionist)**</td>
<td>63.9</td>
</tr>
<tr>
<td>Receptionists worked as Nurse Practitioners as well</td>
<td>13.3</td>
</tr>
<tr>
<td>Nurse Practitioners present in practice</td>
<td>85.6</td>
</tr>
</tbody>
</table>

* Always or often
** Answer options were: yes, no, I do not know

Assigning behaviour

General assigning behaviour

Table 2 shows that the majority of the receptionists intentionally tried to create a varied patient mix (Items 2 and 3) that the majority at least sometimes used patient information to assign (Items 1, 4 and 5) and that the majority seldom used information on the trainee when assigning (Items 6 and 7).

To find out whether the receptionists were aware of a difference in patient mix between trainer and trainee, we asked them to respond to the following statement: ‘the patient mix of the GP trainer and GP trainee do not differ’. Most receptionists agreed or partly agreed [86.3%, 6.5% neutral, 7.1% do (partly) not agree].
Table 2: General assigning behaviour (N=326)

<table>
<thead>
<tr>
<th>I assign the GP trainee:</th>
<th>Often or always (%)</th>
<th>Sometimes (%)</th>
<th>Never or seldom (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Patients who have consulted the GP trainee before</td>
<td>99</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2. Every possible problem</td>
<td>86</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>3. An intentionally varied patient mix</td>
<td>54</td>
<td>21</td>
<td>25</td>
</tr>
<tr>
<td>4. Depends on the reason for the consultation</td>
<td>20</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>5. Relatively easy problems</td>
<td>19</td>
<td>45</td>
<td>36</td>
</tr>
<tr>
<td>6. Patients with an ailment the GP trainee needs to study for the GP Specialty Training</td>
<td>14</td>
<td>30</td>
<td>56</td>
</tr>
<tr>
<td>7. On the basis of my impression of the GP trainee’s qualities</td>
<td>11</td>
<td>26</td>
<td>63</td>
</tr>
</tbody>
</table>

*For reasons of readability we reduced the 5-point Likert scales into 3-point scales.

Assigning subpopulations

In Table 3, the neutral value was predominantly (>50%) selected in nearly all items, indicating that most receptionists had no preference for assigning to the trainee or the trainer. New patients were the only exception: more than half of the receptionists mostly or always assigned them to the GP trainer.

To find patterns in the assigning behaviour with respect to the different subpopulations, we performed a PCA. Five components were identified, resulting in a cumulative explained variance of 58.2%. We interpreted these components as the five assigning patterns shown in Table 3. The first two were complex patients and immigrants. The third pattern consisted of subpopulations with a preference for a specific doctor, a female one or the GP trainer, for instance. The fourth pattern consisted of patients with ‘no preference’ for a specific doctor due to the urgency of their complaint because they were new to the practice or would visit it only once. The fifth pattern consisted of ‘children and minor ailments’.

For ‘complex cases’, >50% of the receptionists reported being neutral in assigning to either trainer or trainee, but 30% reported a preference to assign to the GP–trainer (except for many prescriptions). For ‘minor ailments and children’, the opposite is true: among the minority that was not ‘neutral’, there was a preference for assigning these patients to the trainee. In the ‘no preference’ pattern, this could be distinguished, too, with a clear exception for new patients, as mentioned earlier.
### Table 3. Assigning behaviour for specific subpopulations* (n=326)

<table>
<thead>
<tr>
<th>PCA Component</th>
<th>Explained Variance (%)</th>
<th>Subpopulation</th>
<th>Mostly or always to the GP trainer**</th>
<th>Neutral**</th>
<th>Mostly or always to the GP trainee**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complex</td>
<td>14.0</td>
<td>Psychological problems</td>
<td>32</td>
<td>62</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Difficult cases</td>
<td>38</td>
<td>57</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Many family problems</td>
<td>32</td>
<td>63</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Many prescriptions</td>
<td>3</td>
<td>91</td>
<td>6</td>
</tr>
<tr>
<td>Immigrants</td>
<td>11.5</td>
<td>Immigrants who do speak Dutch</td>
<td>0</td>
<td>96</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Immigrants who do not speak Dutch</td>
<td>2</td>
<td>94</td>
<td>4</td>
</tr>
<tr>
<td>Specific</td>
<td>11.5</td>
<td>Gynaecological problems</td>
<td>5</td>
<td>80</td>
<td>15</td>
</tr>
<tr>
<td>preference</td>
<td></td>
<td>Demanding patients</td>
<td>7</td>
<td>89</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Elderly patients</td>
<td>6</td>
<td>91</td>
<td>4</td>
</tr>
<tr>
<td>No preference</td>
<td>11.3</td>
<td>Emergency cases</td>
<td>18</td>
<td>52</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>New patients</td>
<td>55</td>
<td>37</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Passer-by patients</td>
<td>5</td>
<td>63</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Small surgery</td>
<td>20</td>
<td>54</td>
<td>26</td>
</tr>
<tr>
<td>Children</td>
<td>9.8</td>
<td>Children</td>
<td>0</td>
<td>88</td>
<td>12</td>
</tr>
<tr>
<td>&amp; minor ailments</td>
<td></td>
<td>Minor ailments</td>
<td>0</td>
<td>67</td>
<td>32</td>
</tr>
</tbody>
</table>

* Percentages (abbreviated)
** For reasons of readability we reduced the 5-point Likert scales into 3-point scales.
Discussion

SUMMARY OF THE MAIN FINDINGS
Assigning patients to doctors is the responsibility of medical receptionists. We were interested in their assigning behaviour because it might play an important role in disparities in the patient mix of GP trainees. In our study, we found that almost all receptionists asked patients about their reason for the consultation and the majority at least sometimes discussed which doctor would be seen. In more than three quarters of the practices, no arrangements for assigning during consultation hours were reported and in more than half of the practices, no arrangements for home visits were reported. This suggests that most receptionists have at least some latitude when assigning patients. Receptionists, therefore, play a considerable role in determining a trainee’s patient mix.

However, our results show that most receptionists were not inclined to assign a patient specifically to either the GP trainer or the trainee. They reported intentionally assigning a varied patient mix to trainees. Their strategy in assigning a varied patient mix is not only passive (assigning any patient to any doctor) but active to some extent, too. In general, this can be considered to be a positive result because a varied patient mix that leaves no educational gaps and covers the entire field of primary care is an essential part of a GP’s training.

Receptionists’ assigning behaviour with respect to patient subpopulations shows similar results. The majority of the receptionists did not report a preference when assigning patient groups to GP trainers or trainees.

We found some exceptions. One involved complex patients, who were rarely specifically assigned to trainees, whereas one-third of the receptionists reported having a preference for the GP trainer. For many trainees, handling complex patients is a specific learning objective, mainly in their third training year. Here, a relative underexposure may exist. Minor ailments were preferably assigned to trainees in approximately one-third of cases and new patients were preferably assigned to the GP trainer. The first meeting between the GP and a new patient serves as an opportunity to become mutually acquainted and it therefore makes little sense to assign new patients to a trainee who will reside in the practice for a limited period of time only.

Based on the focus group, we interpreted the subpopulations comprising the pattern ‘specific preference’ for a doctor. We were surprised by the fact that the re-
results suggest that, in general, the receptionists report not assigning these patients to their ‘own’ doctor (the trainer) or a female doctor (often the trainee).

**STRENGTHS AND WEAKNESSES**
As far as we know, this is the first study investigating the assigning behaviour of receptionists. We were, therefore, unable to build on previous research. To increase the content validity of the questionnaire, we used the outcome of a focus group as a basis for construction. This study was based on self-reporting and it cannot be ruled out that receptionists were only partly aware of their actual assigning behaviour.

**OTHER WORK IN THIS AREA**
No other studies addressing the influence of assigning behaviour of medical receptionists on the patient mix of GP trainees have been published, but our findings partly corroborate earlier observational studies about differences in patient mix between trainers and trainees in GP.\(^{31,28}\) Both Eccles and Vintges found that GP trainees saw more acute minor conditions than their GP trainers. Likewise, our study shows that minor ailments were more often assigned to GP trainees.

Eccles and Vintges also found that fewer chronic conditions were seen by trainees. We did not include chronic patients as a separate category, but many are found among elderly patients in combination with the complex pattern. Our results show that complex patients were indeed more frequently assigned to trainers, but elderly patients were assigned neutrally.

Interestingly, Eccles found that ‘female conditions’ were seen by trainees less often, whereas in our study, 80% of the receptionists did not assign these conditions with any preference. Since both studies were conducted at different times and locations, on different scales (62 practices in Northern England and 5 practices in Amsterdam) and with outcome measures different from our study (patient mix as reported in the electronic patient system), it is difficult to draw firm conclusions from these comparisons.

**IMPLICATIONS FOR RESEARCH**
In order to further optimize GP training, new updated observational studies on the composition of patient mix are required. Detecting underexposure in the patient mix can be of value and if an underexposure to complex patients is confirmed, appropriate patients may be steered towards the trainee.
Assigning behaviour is the only factor that can be influenced easily—practice setting is not or hardly changeable and surgery planning can only be influenced to a certain extent. Since our study shows that most receptionists have at least some latitude to steer the patient mix in a specific direction, favourable steering might be accomplished by instructing receptionists. An intervention at the receptionist level seems rational. In a further study, we will therefore investigate whether receptionists’ assigning behaviour can indeed be changed.

If it can, the patient mix of trainees could be adjusted to create better learning and supervision conditions. Better understanding the possibilities of influencing patient mix may also contribute to personalized curricula, tailored to the individual needs of each GP trainee.
REFERENCES


Registrar operating experience over a 15-year period: more, less or more or less the same? Surgeon 2004; 2: 161–4.
Appendix 1.

RESULTS OF THE FOCUS GROUP

One or more of the receptionists in the focus group raised the following issues: They wanted to know the patients’ reason for the consultation and felt that this knowledge influenced their assigning behaviour. They seemed to differentiate between first- and third-year GP trainees: the more experience they presumed a trainee to have, the more complex the problems they tended to assign. They preferred to assign minor ailments to first-year trainees. Other factors that made them assign patients to trainees were patients who had been seen by the trainee before, female patients who preferred a female doctor, emergency cases, patients presenting an ailment the trainee needed to study as part of the specialty training, specific patient groups the trainee explicitly asked to see. They were more likely to assign patients with psychosomatic complaints and other complex social situations to the GP trainer; sometimes, the trainer asked them to do so. They did try to allocate a variety of problems to the surgery hours of the trainee. Some found it difficult to fill the trainee’s surgery hours.
Steering the patient mix of a general practitioner trainee is not as easy as it seems. Results of a randomized controlled intervention.

Jip de Jong, Mechteld R.M. Visser, Margreet Wieringa-de Waard

Provisionally accepted
Abstract

**Background** In studies exploring the patient mixes of general practitioner (GP) trainees, gaps were repeatedly found, as were disparities between the patient mixes of GP trainers and trainees. This reduces the opportunities of trainees to acquire enough competence.

**Objective** The aim of this study was to investigate whether steering the patient mix can be effectuated by instructing medical receptionist, trainer and trainee and to study the effects of this intervention on trainee’s self-efficacy and knowledge.

**Method** Randomized Controlled Trial. After a six-month basic registration period, 73 trainees were randomized. Patients with skin conditions and psychosocial conditions were actively assigned to trainees in the intervention group (n=37) during two successive periods of three months. The patient mix was measured by extracting data from electronic patient records. Learning outcomes were measured by self-efficacy questionnaires and by a knowledge test.

**Results** No increase was found in patient volume and diversity of the steered conditions in the intervention group as compared to the control group. However, the percentual increase of exposure to skin conditions was greater in the intervention group. No difference in skin self-efficacy and psychiatric knowledge was found. The increase of psychosocial self-efficacy was greater in the intervention group. In a regression analysis, patient volume was a significant predictor of both skin and psychosocial self-efficacy.

**Conclusion** Despite the difficulty in implementing steering in daily practice, tailoring the patient mix to the individual learning needs of trainees could be considered.
Introduction

The key factor in learning medical competence is workplace-based clinical exposure\(^1\). Traditionally, medical schools have focused on patient volume: the more patients the trainees see, the better the patient care. This approach has since been replaced by competency-based, outcome-focused training. The traditional approach, however, does have its theoretical merits. The process of developing medical expertise is underpinned by Ericsson’s theory of ‘Deliberate Practice’.\(^2\) Deliberate practice-based curricular planning requires focusing “… on purposefully designed outcome-based learning objectives appropriate content and instructional methods formative assessment feedback reflection and mentoring at each stage of development …”.\(^3\) One can then assume that well-supervised content or ‘patient mix’\(^4\) substantially improves medical competence. Several studies have linked undergraduates’ patient mixes to outcomes.\(^5\)-\(^9\) These have proved inconclusive, although differences in the supervisory quality\(^4\);\(^10\) were found to be more predictive for learning than differences in patient mix.\(^5\);\(^11\)

General practitioner (GP) trainees achieve various competencies and diverse experiences in order to meet required objectives. Consequently, the assumption is that they would most benefit from training when their patient mixes are comparable to those of trainers or later exposed to as licensed physicians.\(^12\)-\(^14\) Problematically, gaps have been found in the patient mix of trainees in different specialty trainings\(^15\)-\(^17\) and, earlier on in medical training,\(^18\)-\(^23\) disparities were found between trainers and their trainees’ patient mixes.\(^12\)-\(^14\);\(^23\);\(^31\) Comparatively, trainees saw more minor ailments,\(^14\);\(^25\);\(^28\) fewer chronic conditions\(^14\);\(^24\);\(^26\);\(^28\) and fewer psychosocial problems.\(^12\);\(^14\);\(^24\);\(^25\);\(^28\) Male trainees were exposed to fewer gynaecological conditions\(^13\);\(^24\);\(^27\);\(^30\) than their female counterparts. Exposure to acute conditions was not univocal.\(^25\);\(^26\);\(^28\)

Several authors have suggested that more balance can be obtained by steering the patient mix in a desired direction.\(^12\)-\(^14\);\(^24\);\(^32\);\(^33\) Steering patients may also occur when an individual trainee needs more exposure to meet a specific learning goal. Steering patients involves practical difficulties because non-urgent patients are free to choose a doctor and do not always wish to consult trainees.\(^34\);\(^35\) Trainee influence on patient mix is likely limited.

Instructing medical receptionists on steering was proposed by Adam and Oswald,\(^32\) an approach requiring awareness of the reason for consultation.\(^33\) Training-practice receptionists often do inquire this of patients.\(^36\)
This study evaluated whether (1) it is possible to intentionally steer trainees’ patient mixes by instructing receptionists, trainers and trainees and (2) to determine whether greater exposure could contribute to better learning. We utilized self-efficacy questionnaires, exposure data and knowledge tests, thereby also addressing the contribution of trainers’ supervisory qualities. We defined ‘patient mix’ by patient volume and diversity.

**Methods**

**Participants**

Forty-nine first-year and 24 third-year trainees from the Institute of GP Specialty Training of the Academic Medical Center, University of Amsterdam participated in this study.

**Study Setting**

We conducted this study between 2008 and 2009 at practices affiliated with GP training where trainees work during their first and third years of a three-year program.

At these sites, trainees perform surgery whilst their trainers are nearby for advice and almost daily coaching sessions. Trainees, trainers and teachers discuss and evaluate learning goals formulated each trimester in addition to assessing progress in clinical competence. Theoretical knowledge is assessed by nationwide testing every 6 months.

**Study Design**

This was a randomized controlled trial. First, we obtained a six-month baseline registration of a trainee’s patient mix, after which two successive interventional trimesters followed. The first steered intervention concerned skin conditions, representing non-complex conditions. During the second trimester, psychiatric, psychological or social (‘psychosocial’) conditions were steered, representing the complex conditions trainees are less often confronted with than their trainers.28

After baseline registration, trainees were randomized into either a control or an intervention group using SPSS 16.0, stratified for training year. To avoid contamination, same-practice trainees were either successively or simultaneously allocated
to the same study arm (n=4).

Trainee self-efficacy (SE) was measured at the start and the end of both intervention periods; knowledge at the start and end of the second trimester. An additional questionnaire was sent to the receptionists of the intervention group to evaluate the steering intervention.

The Ethical Review Board of the Netherlands Association for Medical Education approved this study (NERB-ID 56).

**Intervention**

After randomization, *trainer, trainee and receptionists* at intervention-arm practices were instructed by telephone and letters containing detailed instructions. All three were asked to assign the trainee all patients who presented with skin or psychosocial problems for either that specific consultation or at a more regular basis, or to request these patients to go to the trainee next time. Receptionists received a desk display as a continuous daily reminder and a laminated card listing the interventional conditions and symptoms in colloquial language. The trainer and trainee received a laminated card listing the interventional ICPC codes. Understanding and progression of the intervention were assessed by phone three weeks after start.

**INSTRUMENTS**

**Patient mix**

In the Netherlands, most GPs keep detailed Electronic Patient Records (EPRs) which include a standardized system of diagnostic codes: the International Classification of Primary Care (ICPC). The ICPC consists of 17 organ system-related chapters, a general chapter and a social chapter. Data were extracted from EPRs at the end of each trimester; as these were aggregated, information on individual patients was not retrievable.

**Self-efficacy**

Based on Bandura’s methodology, we developed two 20-item SE-questionnaires on skin and psychosocial conditions (Text box 1). Both questionnaires contained items relating to general, diagnostic and therapeutic competencies concerning specific conditions, a representative sample of the entire spectrum of skin and psycho-
social diagnoses originating from relevant ICPC chapters. Questionnaire content was discussed by the study advisory committee and tested by six experienced GPs after which they were finalized.

Trainees could answer the questions on a visual analogue scale ranging from ‘not capable’ to ‘fully capable’. Scales were scored using a scale ruler overlay ranging from 0 to 100.

**Knowledge test**

To address the effect of the intervention on a more formal, objective learning outcome, an online knowledge test was taken on one of the sub-domains of the intervention. Knowledge about psychiatric (not social) conditions relevant to daily patient care was assessed. This test was developed by the Dutch GP Training network. To correct for potential differences in difficulty between the pre-intervention and the post-intervention knowledge test and to avoid a memory effect, two versions (A and B) were composed, each containing 50 patient case-based items of the true-false-type and an additional ‘I don’t know’ option, covering the entire spectrum of primary-care psychiatry. Half the (randomly assigned) group received version A as a pre-intervention knowledge test and version B post-intervention; the other half, the opposite. Trainees could take the test online restricted to 30 minutes. The test was scored as correct minus incorrect answers. The tests were used exclusively for research purposes.

**Quality of supervision**

At the end of the study, trainees evaluated the quality of trainers’ supervision by completing the Dutch translation of the validated Cleveland Clinic’s Teaching Effectiveness Inventory (CCTEI), which contains 15 variables on trainer quality. Possible answers ranged from never/poor (0) to always/superb (5); averages are calculated.

**Questionnaire for medical receptionists**

This concerned the perceived ability to steer relevant patients to trainees during both intervention periods.
**Text box 1. Examples of items on the self-efficacy questionnaire**

The following questions pertain to tasks performed by the general practitioner only. Please indicate how well you find yourself capable of:

<table>
<thead>
<tr>
<th>Self-efficacy in skin conditions</th>
<th>Not capable</th>
<th>Fully capable</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Diagnosing skin conditions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Local treatment of skin conditions</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Specific</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Recognizing erysipelas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Treating erysipelas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Treating contact dermatitis</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Self-efficacy in psychosocial conditions</th>
<th>Not capable</th>
<th>Fully capable</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Diagnosing psychological/psychiatric conditions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Recognizing social problems</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Specific</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Recognizing depression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Treating depression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Treating alcohol addiction</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
STATISTICAL ANALYSES

**Patient mix**
Participants were excluded from the patient mix analyses if less than 50% of their patient contacts had an ICPC diagnosis code. We analyzed two components of patient mix, ICPC diagnosis volume (patient volume) and ICPC diagnosis diversity (diversity). Percentages of skin/psychosocial conditions compared with all encountered diagnoses were calculated by setting the total of all ICPC diagnosis codes to 100%, disregarding those without an ICPC diagnosis code. For both absolute numbers and percentages, means of available baseline-registration trimesters were established.

Diversity of patient mix was expressed as the number of different ICPC diagnosis codes each trainee had encountered. Differences between intervention and control groups were tested using variance analysis for repeated measures (ANOVA, SPSS 18.0); excluding separate first- and third-year subgroup analyses (their numbers would have been insufficient).

**Self-efficacy and knowledge tests**
Internal consistency was calculated for each of the four SE-questionnaires by using Cronbach’s alpha. The 20-item scores were averaged for each participant. Differences in increase in SE and knowledge between the intervention and control group were established using ANOVA. Effect sizes were calculated using Cohen’s d.

**Supervision**
For the CCTEI questionnaire, Cronbach’s alpha was computed; average score of the constituting items was then used. Whether to adjust for supervision quality depended on the existence of significant differences between the arms.

**Intervention evaluation questionnaire for receptionists**
We performed ANOVA to find relationships between the ability to steer patients and increases in patients actually seen. Independent variables: frequency of eligible patients encountered by receptionists, how many had been assigned to trainees, most successful steering period. Dependent variables: absolute and proportional increases and actual, intervention-trimester exposure to skin or psychosocial conditions. Scheffe’s post-hoc testing ensued if significant effects were found.
Results

PARTICIPANTS AND FOLLOW UP
Seventy-three practices were included (Figure 1) of which 49 accommodated first-year and 24 accommodated third-year trainees; 19 men and 54 women starting a new training period. All gave informed consent. The first-year trainees had started in September of 2008; third-year trainees in March, June, September or December of 2008. Sixty-two trainees had started a 12-month practical period; eleven trainees 9 months. First-year trainees of the latter group had completed the 6-month baseline registration but only took part in the first intervention period, whereas the third-year trainees took part in both intervention periods with their baseline-registration period being shortened to three months.

Two trainees’ ICPC-data were disregarded because of mean coding percentages below 50%. Six trainer-trainee pairs were nullified during the study; the trainees were moved to other practices. Four of them (1 control, 3 intervention) had unsuitable EPRs; the trainees’ extraction data could not be retrieved although one trainee could later be re-included. The remaining two trainees continued by instructing the receptionists and the trainers on their new practices. All six trainees had returned their SE-Questionnaires/knowledge tests.
Note that completion of the questionnaires and knowledge tests was conducted separately from EPR data collection.
PATIENT VOLUME
Percentual increase of skin conditions differed significantly between trial arms (Table 1), this was not found for psychosocial conditions.

Table I. Exposure to skin and psychosocial conditions (numbers and percentages) for the control and intervention groups, test results and effect sizes

<table>
<thead>
<tr>
<th></th>
<th>Control (n=34) Mean (SD)</th>
<th>Intervention (n=35/30*) Mean (SD)</th>
<th>Effect Sizea</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Abs</td>
<td>%</td>
<td>Abs</td>
</tr>
<tr>
<td>Skin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline registrationb</td>
<td>90.3 (29.3)</td>
<td>16.9 (3.3)</td>
<td>88.4 (28.5)</td>
</tr>
<tr>
<td>Intervention trimester</td>
<td>104.0 (34.2)</td>
<td>16.5 (2.9)</td>
<td>111.1 (39.9)</td>
</tr>
<tr>
<td>Increase</td>
<td>13.7 (33.3)</td>
<td>-0.4 (3.3)</td>
<td>22.7 (33.7)</td>
</tr>
<tr>
<td>Psychosocial</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline registrationb</td>
<td>25.8 (14.5)</td>
<td>4.7 (2.1)</td>
<td>28.6 (11.8)</td>
</tr>
<tr>
<td>Intervention trimester</td>
<td>33.0 (18.3)</td>
<td>5.1 (1.8)</td>
<td>34.3 (15.2)</td>
</tr>
<tr>
<td>Increase</td>
<td>7.3 (16.0)</td>
<td>0.4 (1.5)</td>
<td>5.8 (12.9)</td>
</tr>
</tbody>
</table>

* See flow chart
† p<0.05
a Increase Intervention - Increase controls / (√((SD Increase Intervention^2 + SD Increase Controls^2) /2))
b Mean of two trimesters

DIVERSITY
No significant differences (p>0.05) were found between the intervention and the control groups in diversity of skin diagnoses (m(int)_n=30.6, SD=5.4.; m(control)_n=31.3, SD=5.3, range 20-42 of 58 possible ICPC-codes), nor in diversity of psychosocial conditions seen (m(int)=14.6, SD =4.3; m(control)=13.1, SD=4.3, range 3-21 of 68 possible ICPC-codes).

QUALITY OF SUPERVISION
Average CCTEI-score was 4.0 (SD=0.6). Cronbach’s alpha was 0.90. Since no significant difference was found between the control (4.0, SD=0.5) and the intervention groups (3.9, SD=0.6), no correction for supervision in the analyses took place.
Chapter 7

Self-Efficacy

Cronbach’s alpha of the SE-skin questionnaire was 0.88, and 0.94 for the 20 SE-psychosocial conditions items, pre- and post-intervention. Intervention/control group-increases in SE (Table 2) were significant for both skin and psychosocial conditions (p<0.05); the increase for psychosocial conditions being greater in the intervention group (P<0.05).

Table 2. Overall self-efficacy, pre- and post-intervention; mean (SD) of 20 items

<table>
<thead>
<tr>
<th></th>
<th>Control (Mean (SD)) (n=34/33*)</th>
<th>Intervention (Mean (SD)) (n=36/34*)</th>
<th>Effect Sizea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-SE</td>
<td>69.7 (9.4)</td>
<td>65.9 (9.2)</td>
<td></td>
</tr>
<tr>
<td>Post-SE</td>
<td>75.5 (8.6)</td>
<td>73.2 (7.7)</td>
<td></td>
</tr>
<tr>
<td>Increase</td>
<td>5.8 (9.3)</td>
<td>7.3 (7.8)</td>
<td>0.17</td>
</tr>
<tr>
<td>Psychosocial</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-SE</td>
<td>66.9 (10.9)</td>
<td>57.9 (10.4)</td>
<td></td>
</tr>
<tr>
<td>Post-SE</td>
<td>66.6 (9.9)</td>
<td>63.7† (10.0)</td>
<td></td>
</tr>
<tr>
<td>Increase</td>
<td>1.8 (9.2)</td>
<td>5.8 (6.5)</td>
<td>0.50</td>
</tr>
</tbody>
</table>

*a Cases (trainees) analysed. Skin/Psychosocial
† p<0.05
a (Increase Intervention - Increase controls / (sqrt((SD Increase Intervention^2 + SD Increase Controls^2) /2)))

Knowledge of Psychiatric Conditions

Knowledge increased in the control group (m=1.5; SD=9.8) and decreased (m=-1.9; SD=8.0) in the intervention group, p<0.05. Version A was more difficult than B as scores were lower: p<0.05 (see Appendix, Supplementary Table 1). Knowledge increase did not differ between the arms, if corrected for A-or-B-version (p>0.05).

Responses to Receptionists’ Questionnaire (Table 3)

All receptionists in the intervention arm (n=35) responded; however, two from one practice (their scores were averaged and analyzed as one subject). Analyses and post-hoc tests showed relationships (p<0.05) between those eligible patients having skin conditions assigned to trainees (Table 3, item 6) and the increase in/actual percentages of patient volume. A significant relationship existed between the most successful allocation period (item 7) and actual absolute volume (p<0.05). Comparisons for psychosocial conditions were not significant.
Table 3. Questionnaire for the medical receptionists (n=35) of the intervention group

<table>
<thead>
<tr>
<th>Items</th>
<th>Scales, Anchors and Scores</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. How often do you ask a patient the reason for consultation?</td>
<td>Always % (n)</td>
</tr>
<tr>
<td></td>
<td>59 (20)</td>
</tr>
<tr>
<td>2. Did you find the letter with information about the intervention clear?</td>
<td>Very clear</td>
</tr>
<tr>
<td></td>
<td>60 (21)</td>
</tr>
<tr>
<td>3. Did you find the desk display a helpful reminder for the intervention?</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>27 (9)</td>
</tr>
<tr>
<td>4. Did you find the reference cards helpful for remembering the intervention?</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>56 (19)</td>
</tr>
<tr>
<td>5. How frequently did you encounter an eligible patient?</td>
<td>≥ Once a day</td>
</tr>
<tr>
<td>Skin</td>
<td>17 (5)</td>
</tr>
<tr>
<td>Psychosocial</td>
<td>21 (6)</td>
</tr>
<tr>
<td>6. Which part of the eligible-patient group did you assign to the trainee?</td>
<td>All patients</td>
</tr>
<tr>
<td>Skin*</td>
<td>19 (6)</td>
</tr>
<tr>
<td>Psychosocial</td>
<td>-</td>
</tr>
<tr>
<td>7. Which period of interventional steering was the easiest?</td>
<td>Mainly the start</td>
</tr>
<tr>
<td>Skin conditions*</td>
<td>3 (1)</td>
</tr>
<tr>
<td>Psychosocial</td>
<td>-</td>
</tr>
</tbody>
</table>

*p<0.05; see text.

RELATIONSHIP BETWEEN PATIENT MIX AND LEARNING

Because the steering intervention showed no univocal increase in patient mix, the relationship between patient mix and learning was explored further in regression analyses (method=enter). Trainees from control and intervention groups were combined. Dependent variables were post-SE questionnaires and post-intervention-knowledge-test scores (see Appendix, Supplementary Table 2). Predictors were entered first on which the intervention could have had no effect (pre-SE ques-
tionnaires/knowledge test, CCTEI score, knowledge A-or-B-version and intervention). Patient mix was then added in a second block. High Pearson correlations between volume and diversity (0.72-0.87) indicated co-linearity. Therefore, we did not enter diversity as a separate predictor, as volume had been the primary aim of the steering interventions.

The pre-SE/knowledge test and the A-or-B-version were found to be significant predictors (p<0.05). Volume predicted both SE-skin and SE-psychosocial (p<0.05); explained variance increased from 29% to 34% for post-SE skin, from 53% to 58% for post-SE psychosocial, and remained at 28% for post-psychiatric knowledge.

**Discussion**

We conducted a randomized controlled trial to investigate whether a trainee’s patient mix could be actively steered and the effects of this on learning. We carefully concluded that steering of skin conditions was possible. Although in absolute number only a non-significant trend could be observed, there was a higher percentual increase in the intervention-control comparison; however not leading to greater SE. Comparatively, psychosocial-condition-volume did not increase but greater intervention group-SE was identified. Knowledge of psychiatric conditions neither increased between pre- and post-testing, nor between the arms. Post-SE was predicted by pre-test SE and by volume. Quality of supervision was equal in both arms so we did not adjust for it. Supervision neither predicted SE alterations nor knowledge.

The lack of increase in SE-skin, despite slightly successful steering, may be due to skin conditions already being frequently encountered.\(^{27}\) Consequently, additional cases would not contribute much to learning because the flattened part of the competence growth curve would have been reached.\(^{42;43}\) Moreover, we found a percentual increase in volume, but not in diversity.

As to the increase in self-efficacy in complex (psychosocial) conditions, we conjectured that being allocated to an intervention may have increased trainee ‘awareness’ and therefore more learning activities and more ‘deliberate practice’. Perhaps they explored and discussed such cases more than they otherwise would have.

The limited success of the intervention was contrary to the intervention group receptionists’ perceptions. Most receptionists ‘more than regularly’ asked patients
the reason for their consultations when planning the trainees’ surgeries, which was an important condition for successful steering. Questionnaire results clearly show that they supposed they were steering substantial numbers of skin conditions. Receptions found steering psychosocial conditions more difficult; results suggest however, that also in this category, substantial steering occurred.

Although we instructed both trainers and trainees, they had little opportunity to steer patients themselves. Stimulation by instituting target values, for instance, might have been helpful. Effectiveness might have increased if we had sent more frequent reminders to them.

COMPARISON WITH OTHER STUDIES

Most studies exploring the relationship between patient mix and learning were performed with medical students or residents. The quality of designs varied and, to our knowledge, there have been no randomized trials on this. Some studies showed relationships between patient mix and self-reported learning like the positive relation between volume and SE we found. This result was congruent with findings by Hoifoidt in that more experience was related to greater subjective learning during a psychiatric rotation. Dolmans found that patient mix and supervision both positively influenced a rotation’s effectiveness, but supervision was more essential because high-quality supervision guaranteed sufficiently scored rotation effectiveness, irrespective of patient mix. Despite other studies’ emphasis on supervision, it was not predictive of learning in our study. One could question whether the importance of supervision also applies to long-term medical training as in our one-on-one settings. Here, the trainers themselves are solidly trained in didactic competencies. One could speculate that the contribution of supervision to learning decreases over time, since its significance was mainly found in relatively short medical student-rotations supervised by numerous preceptors.

It is surprising that studies predicting performance from patient exposure have often failed to show a relationship between exposure and a more objective learning outcome, such as OSCE’s, clinical assessment or multiple choice or other written examinations, like the knowledge test we used. For instance, Jolly found no relationship between clinical experience and educational outcomes and Greenberg found no impact of clinical experience on a paediatric subject test. An exception to this was a study by Lampe, who found that the requirement of 10 predetermined cases to be exposed to improved knowledge and performance
in an emergency medicine rotation. For this, future studies might consider the combination of steering instructions with prerequisite exposure numbers. A rational consensus must first be reached about the numbers required. Fung questioned the feasibility of full implementation of deliberate practice, after a correlation study showed no relationship between exposure and OSCE score. This was attributed to a lack of clear learning goals, quality and quantity of feedback and reflection. In our setting, these conditions were generally well-organized; it remains obscure why the relation patient mix-learning appears intangible.

**STRENGTHS AND LIMITATIONS**

A strength of our study was the use of data from EPR systems to describe patient mix. Many studies used handwritten or digital logbooks intended solely for educational purposes. This type of data is subject to social-desirability bias by the trainee. Other strengths were RCT design, use of both subjective and objective learning outcome measures, and comparison to baseline registration. Ericsson found that effective practice is characterized by clearly focused goals, time-commitment, immediate feedback, and reflection. In our setting, trainees are compelled to formulate learning goals; time/feedback opportunity is generally sufficient as was demonstrated by CCTEI high mean scores. The setting was therefore appropriate for specifically studying a patient mix-contribution to learning.

A major limitation was that the intervention had not been tailored to individual trainee learning needs. Such an approach would require the development of multiple learning outcome instruments not considered feasible for this study. This untailored approach may have hampered trainer-trainee motivation to steer. Furthermore, while we did measure trainers’ global supervisory qualities, we did not do so for steered conditions specifically.

The one-trimester duration of each intervention was probably too short for major effects. Since a (non-significant) trend towards higher volume of skin diseases was observed, extending the steering period might have revealed clearer results. It remains questionable, though, how long receptionists could continue steering.
SUGGESTIONS FOR FURTHER RESEARCH

This was the first RCT on the effect of active patient steering in practical GP specialty training. We found that non-complex conditions could be steered whereas complex conditions could not; surprisingly, learning effects (SE) revealed the opposite. Like other studies on medical training, we succeeded in linking patient mix (volume) to self-reported learning, but not to objective learning.

Future studies should explore the effects of interventions, tailoring patient mix to individual learning aims so as to achieve better implementation of deliberate-practice-based curricular planning. Instructing receptionists to steer could achieve this aim. The contribution of predefined prerequisite exposure numbers, as formulated by trainees themselves or expert consensus should be further explored. Despite the crucial position of supervision in undergraduate settings, its specific contribution in longer-lasting training relationships needs further exploration.
REFERENCES

20. Raghoobar-Krieger HM, Sleijfer DT, Hofs-tee WK, Kreeftenberg HG, Bender W. The


40. Busari JO, Weggelaar NM, Knottnerus AC,
151


42. ten Cate T, Snell L, Carraccio C. Medical competence: The interplay between individual ability and the health care environment. Med Teach 2010; 32(8):669-675.

43. Albanese M. Problem-based learning: why curricula are likely to show little effect on knowledge and clinical skills. Med Educ 2000; 34(9):729-738.


# Appendix 1

## Supplementary Table 1.

**Knowledge of psychiatric conditions**

<table>
<thead>
<tr>
<th></th>
<th>Control Mean (SD) n=33</th>
<th>Intervention Mean (SD) n=34</th>
<th>Version A-B n=33</th>
<th>Version B-A n=34</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-knowledge</td>
<td>25.7 (7.3)</td>
<td>28.9 (7.3)</td>
<td>25.2 (7.8)</td>
<td>26.4 (6.5)</td>
</tr>
<tr>
<td>Post-knowledge</td>
<td>27.2 (7.3)</td>
<td>27.0 (7.1)</td>
<td>29.9 (6.1)</td>
<td>24.4 (7.1)</td>
</tr>
<tr>
<td>Increase</td>
<td>1.5 (9.8)</td>
<td>-1.9 (8.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Effect Size</strong></td>
<td></td>
<td></td>
<td>0.38</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Control Mean (SD) n=15</th>
<th>Intervention Mean (SD) n=18</th>
<th>Version B-A n=18</th>
<th>Control Mean (SD) n=18</th>
<th>Intervention Mean (SD) n=16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-knowledge</td>
<td>23.1 (7.0)</td>
<td>26.9 (8.2)</td>
<td>Pre-knowledge</td>
<td>27.8 (7.0)</td>
<td>31.2 (5.5)</td>
</tr>
<tr>
<td>Post-knowledge</td>
<td>30.5 (6.5)</td>
<td>29.4 (5.9)</td>
<td>Post-knowledge</td>
<td>24.5 (7.0)</td>
<td>24.2 (7.4)</td>
</tr>
<tr>
<td>Increase</td>
<td>7.4 (5.8)</td>
<td>2.5 (7.0)</td>
<td>Increase</td>
<td>-3.3 (9.9)</td>
<td>-7.0 (5.8)</td>
</tr>
<tr>
<td><strong>Effect Size</strong></td>
<td></td>
<td></td>
<td></td>
<td>0.76</td>
<td>0.46</td>
</tr>
</tbody>
</table>

* Increase Intervention - Increase controls/ (√((SD Increase Intervention² + SD Increase Controls²) /2))
Supplementary Table 2.
Regression analyses of factors determining learning outcome

<table>
<thead>
<tr>
<th></th>
<th>Post-SE Skin Conditions</th>
<th>Post-SE Psychosocial Conditions</th>
<th>Post- Psychiatric Knowledge</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R² Block1 =0.29</td>
<td>R² Block1 =0.53</td>
<td>R² Block1 =0.29</td>
</tr>
<tr>
<td></td>
<td>R² Block2 =0.34</td>
<td>R² Block2 =0.58</td>
<td>R² Block2 =0.28</td>
</tr>
<tr>
<td></td>
<td>Beta (p)</td>
<td>Beta (p)</td>
<td>Beta (p)</td>
</tr>
<tr>
<td></td>
<td>n=65</td>
<td>n=61</td>
<td>n=61</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Block 1</th>
<th>Pre-SE</th>
<th>Intervention</th>
<th>Mean CCTEI a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.53 (0.00)*</td>
<td>-0.34 (0.76)</td>
<td>0.07 (0.54)</td>
</tr>
<tr>
<td></td>
<td>0.73 (0.00)*</td>
<td>0.01 (0.94)</td>
<td>0.00 (0.99)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.16 (0.15)</td>
<td>-0.12 (0.29)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.34 (0.01)*</td>
<td>-0.54 (0.00)*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Block 2</th>
<th>Pre-SE</th>
<th>Intervention</th>
<th>Mean CCTEI a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.50 (0.00)*</td>
<td>-0.06 (0.61)</td>
<td>0.06 (0.59)</td>
</tr>
<tr>
<td></td>
<td>0.68 (0.00)*</td>
<td>-0.28 (0.77)</td>
<td>-0.00 (0.99)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-0.17 (0.15)</td>
<td>0.33 (0.01)*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-0.55 (0.00)*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Patient Mix Volume b</td>
<td>0.21 (0.05)*</td>
<td>0.22 (0.02)*</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.07 (0.57)</td>
</tr>
</tbody>
</table>

* Denotes significance

a Cleveland Clinic’s Teaching Effectiveness Inventory

b For post-psychiatric knowledge, this was the patient mix involving psychiatric conditions only. For the post SE-psychosocial conditions, the patient mixes involving both psychiatric and social conditions were measured.
CHAPTER 8

General Discussion
General Discussion

1. Introduction to the general discussion

This thesis was undertaken to gain a better insight into the patient mix that GP trainees are confronted with during their specialty training. To this end, a cohort study was started which enrolled 73 GP trainees and their trainers. The patient mixes of trainees and trainers were compared and areas of low exposure were identified. Reasons for the differences in the patient mixes of trainers and trainees due to the assigning behaviour of receptionists were addressed. The cohort study was followed by a randomized controlled trial (RCT) in which the patient mix of half of the trainees was actively steered on skin diseases and psychosocial conditions successively. The aim of the RCT was to study whether steering was possible and, if so, whether it would lead to a better learning outcome. Learning outcome was measured by self-efficacy questionnaires and a knowledge test. Because our instrument for describing patient mix was based on the ICPC coding, the barriers that the GP trainees and trainers faced in their daily clinical work to ICPC coding and specific situations prone to non-coding were studied. To obtain more insight into the relationship between patient mix and learning, a systematic review of this relationship in work-based clinical settings was carried out.

In the first part of this general discussion, the main findings are summarized per chapter and the interpretation and implications are discussed. The results of Chapters 4 and 5 are discussed together. In the second part, more general educational implications for the GP specialty training are discussed. Finally some recommendations for future research are given.

2. Main findings, interpretation, and implications

**CHAPTER 2: WHAT IS KNOWN ABOUT THE RELATIONSHIP BETWEEN PATIENT MIX AND LEARNING IN WORK-BASED CLINICAL SETTINGS?**

**Main findings:** In Chapter 2, the evidence of the relationship between patient mix and learning in the field of medical education was systematically reviewed. Descriptions of patient mix in the 22 reviewed studies were heterogeneous. They were related to different outcomes in different settings, at different stages of education.
Patient mix was found to be related to self-reported learning outcome, the experienced quality of the training programme being most evident. Other learning-related factors, such as supervision and learning style, seemed to mediate the relationship between patient mix and learning. A relationship between patient mix and formal assessment was rarely demonstrated.

**Interpretation and implications:** Interpretation of the evidence was hampered by the absence of a widely accepted definition of the concept of ‘patient mix’. The patient mix as presented often seemed to depend on the instrument the authors had at their disposal more than on a specific operationalization of the patient mix under investigation. The interventions on the patient mix that we found were indirect (due to curriculum change); not meant to actively influence the patient mix.

It can be questioned whether formal assessments are sensitive enough to ascertain how patient mix can affect learning outcome. The positive relationship between patient mix and self-assessment outcome is not per se an indication of a relationship between patient mix and learning. The absence of a relationship between patient mix and formal assessments might lead to the conclusion that increasing experience does not *automatically* lead to more competence. This is in line with the deliberate practice theory (see Introduction), which states that feedback is required. The relationship between patient mix and learning is clearly complex, and many other variables play a role (such as supervision quality, learning style, learning environment and professionalism). Educational research would benefit from a standardized approach in patient mix descriptions. An inquiry into more detailed aspects of patient mix in relation to their contribution to learning is desirable.

**CHAPTER 3: WHAT IS THE RELATIONSHIP BETWEEN THE BARRIERS TO ICPC CODING OF GP TRAINEES AND TRAINERS AND THEIR SELF-REPORTED AND ACTUAL CODING PERFORMANCE?**

**Main findings:** To gain insight into the validity of our patient mix descriptions, the coding behaviour of the trainers and trainees was investigated in a questionnaire. Results were compared with coding performance as measured by EPR extractions. Hence, coding bias, due to the barriers that GP trainers and trainees experience while using the ICPC classification system, was addressed.

The percentages of contacts with an ICPC code were high, both for trainers and for trainees. They reported that most barriers were of minor importance. Coding performance was best predicted by motivation- and ICPC-related barriers.
Trainers’ estimation of their coding performance was reasonably accurate, considering the high correlation with the actual coding percentage. For trainees, this correlation was lower. Trainers’ ICPC coding behaviour seemed to be influenced more by the extent to which they experience barriers than the trainees’ coding behaviour.

Interpretation and implications: The only situations we found to be prone to non-coding were telephone consultations, repeat prescriptions, and administrative actions. The latter two are of minor importance since they were disregarded in our patient mix descriptions. EPR-based data thus seem to provide a representative reflection of the patient mix seen by trainers and trainees. This makes these data suitable for research and educational purposes, as used in this thesis. To further reduce the chance of biased results, the content of the non-coded telephone consultations should be addressed in future studies.

CHAPTER 4 & 5: THE PATIENT MIX THAT GP TRAINEES ENCOUNTER IN GP TRAINING PRACTICE: DIFFERENCES BETWEEN FIRST- AND THIRD-YEAR TRAINEES AND BETWEEN TRAINEES AND THEIR TRAINERS, IN BOTH HIGH-AND LOW-EXPOSURE AREAS

In the following paragraphs, the relevance of the most meaningful disparities that were found is discussed. Implications of these disparities and areas of low exposure are considered in relation to the health care that GPs are required to provide as described by the Dutch College of General Practitioners (NHG) and the Dutch Association of General Practitioners (LHV) in ‘Health Care Supply General Practice 2009’ and in relation to the ‘Competency Profile and Final Requirements of the General Practitioner’ by the Dutch GP Training Institutes. In addition, implications due to demographic trends are commented on, led by the NHG and LHV’s ‘Vision for the future 2012’ which will be updated shortly.

Contact numbers

Main findings: The mean number of face-to-face consultations per trimester did not significantly differ between the first and third year of training (see Chapter 4), but the standard deviation almost doubled in the third year, indicating more variance in the number of patient contacts among third-year trainees. Closer examination showed that this was largely due to several trainees who saw exceptionally large numbers of patients. Trainers did about double the number of consultations and home visits, and treble the number of telephone consultations compared with train-
This cannot be fully explained by their working hours.

**Interpretation and implications:** Trainees who see large numbers of patients build more experience. To what extent this experience also will lead to more learning and, ultimately, better performance is unknown. It would be interesting to study which part of their higher workload is due to individual characteristics of these trainees, such as their working speed, their trainers’ characteristics, or the organizational factors of the training practice, such as the practice setting and surgery planning (see Chapter 5).

The finding that GP trainees are exposed to half the amount of patients compared with their trainers, poses the question whether the trainees are sufficiently prepared to handle the workload when starting to work as an autonomous GP. This emphasizes the importance of the ‘autonomous’ periods in which the trainee is working alone in the practice without the presence of their supervisor. In these periods, they become better prepared to the working speed of regular GP surgery hours. Interventions to increase the work speed, e.g. extension of the autonomous period, could help to prepare trainees for their future job. The workload of a regular GP is a challenge for newly qualified GPs when running a practice themselves. The ability to keep control on the higher workload is perhaps one of the reasons why they have a preference for working as a locum doctor or under contract, instead of immediately starting a practice themselves.\(^7\)

The total number of patient contacts per trimester in training practices is about the same as in regular (non-training) GP practices.\(^8\,^9\) We did not investigate whether there are other differences in patient populations between regular and training practices. Comparison of our data with national data (Chapter 5), did not point in that direction.

**Age distribution**

**Main findings:** Significant differences between trainers and trainees were found in all patient age groups, except the 25–44 and 75+ year groups. In general, trainees see a higher proportion of younger patients, whereas the opposite is the case for their trainers.

**Interpretation and implications:** The reasons for this uneven age distribution may well lie in other patient mix characteristics which relate to age: trainees see fewer chronic conditions, which are more often present in older patients. Patients with chronic conditions may feel more comfortable with the trainer.\(^10\) In addition,
trainees see more minor ailments and respiratory infections, which are more often present in children. The latter can be regarded as positive, because becoming familiar with care for children is an important learning objective for many trainees. Furthermore, by seeing children together with their parent(s) a trainee becomes familiar with the family doctor’s role.

**Chronic diseases**

**Main findings:** The exposure to chronic diseases increased steadily during the training. Trainees saw a smaller proportion of chronic conditions than their trainers (first year: 8.7% vs. 15.8% and third year: 10.8% vs. 16.2%, Chapter 5). Similar results were found for oncological conditions (which come under the category of chronic conditions; 0.4% vs. 1.7% and 0.9% vs. 1.8%).

**Interpretation and implications:** In Chapter 4, it is explained that the actual number of chronic conditions encountered by trainees (and trainers) is probably higher than our findings suggest, as our software was programmed to count actual health problems and not co-morbid chronic conditions. The steadily increase in exposure is in line with the learning objectives, as the third-year curriculum places emphasis on chronic and complex conditions. However, as is shown in Chapter 5, in the third year, a substantial disparity still was found between the proportional exposure of trainers and trainees. Our results thus confirm the finding of earlier studies that GP trainees saw fewer chronic conditions than trainers. Several authors have warned about under-exposure of trainees to chronic conditions. Darer et al. found that the majority of US GPs felt inadequately trained for chronic diseases, and Card et al. reported that recently graduated internists felt insufficiently prepared for the care of patients with chronic conditions. Given the challenges of the ageing population, it can be concluded that trainees should handle more patients with chronic disease. This is also in agreement with recommendations in the ‘position papers’ of the NHG regarding chronic disease in which the importance of preparing GP trainees for care of the elderly and for complex care is emphasized, and cooperation between GP training programmes and other primary (nursing home doctors) and secondary care specialists (geriatricians, specialists in geriatric psychiatry) is recommended.

Trainees also saw very few oncological conditions. The recommendations in a report on the GP’s role in the aftercare of cancer patients, which was recently published by the Dutch Cancer Society, pleaded for a much larger role of the GP in the
aftercare and the organization of care for the oncological patient. In this report, tuning and synchronizing the care provided by the GP and the medical specialist is emphasized. The committee recommends that combined training programmes are set up for primary care specialists (including GPs) and secondary care specialists informing them about how to act in the different phases of the oncological chain of care. Similar to chronic conditions, oncological conditions may have been under-reported, but the level of exposure to oncological conditions we found in Chapters 4 and 5 is likely insufficient to fulfil these training recommendations. The low exposure might change in the future due to the ageing population, but until then, the second-year ‘chronic disease’ rotation in nursing homes, clinical geriatric wards or in palliative care, maintains an important part of the current GP training programme.

Female conditions

Main findings: In line with other studies of medical students and family medicine residents, female conditions were seen more often by female trainees than by male trainees (Chapter 4). Our analyses in Chapter 5 showed that differences in seeing female conditions were primarily related to the doctor’s gender.

Interpretation and implications: Eccles et al., Fleming and others found that male and female trainees saw fewer female conditions than their trainers did whereas our results showed that this was primarily related to the gender of the doctor. As the majority of the GP trainees nowadays are female, perhaps a former preference for the more experienced trainer (predominantly male) is nowadays counteracted by a stronger preference for a female doctor. Assuming that male and female GPs should be equally competent, our findings can have implications for the competence development of male trainees. Whether the males reach the required level of competence should be further studied.

Male conditions

Main findings: The numbers of male conditions encountered were low and did not significantly differ between female and male trainees (Chapter 4). However, male doctors (trainees and trainers combined) encountered a significant larger proportion of male conditions than female doctors (Chapter 5).

Interpretation and implications: The low numbers of male conditions questions whether there were sufficient encounters for trainees to acquire competence. A potential, gender-related, underexposure was also addressed by Levy and Merchant,
who found that male students received more experience with male-specific examination skills. However, since the numbers and percentages we found were low, one could question whether the significant difference is also educationally relevant.

**Psychological and psychiatric conditions**

**Main findings:** Similar to the results in other studies,\textsuperscript{12,25,26,28,31} the trainees saw fewer psychiatric conditions than their trainers (Chapter 5). This was more pronounced in first-year than in third-year trainees (Chapter 4). Compared with national morbidity data,\textsuperscript{32} the difference in exposure between trainees and regular GPs was less.

**Interpretation and implications:** An explanation for the differences found may be that one third of the medical receptionists tend to assign patients with psychological problems to the trainer (Chapter 6, see also next paragraph). The ‘position paper’ of the NHG on Psychiatric care in General Practice is explicit about the central role and far-reaching responsibilities the GP should have in the care of patients with acute and chronic psychological problems.\textsuperscript{33} This states, among other things, that the GP is responsible for diagnosing psychiatric problems, orchestrating the care process, identifying patients (including children) at risk of developing psychiatric problems, taking preventive interventions, and caring for psychiatric patients who cannot be treated or can no longer be treated in secondary care. Training in preventive tasks is explicitly recommended, as is cooperation with other professionals. It is questionable whether the actual exposure provides enough learning opportunity to fulfil the demands of this ambitious future vision.

Our intervention intended to increase the exposure to psychiatric and social conditions was not successful (Chapter 7, see further). However, seeing more patients with these conditions did improve self-efficacy in this area. This makes it worthwhile to consider stronger interventions to steer the patient mix in this area. Additional measures, such as requiring pre-specified contact numbers, can be helpful to accomplish this. In addition, the rotation in the second year of the training in inpatient and outpatient psychiatric clinics maintains an important addition to the training programme.

**Social conditions**

**Main findings:** The exposure to social problems seemed low, and the exposure was even lower for the trainees than for the trainers. Trainers with a coding percentage
Unlimited Exposure

over 90% saw more psychiatry and social problems than trainers within the 50–90% range, indicating that psychosocial codes are under-reported by some trainers. This was not found to be the case for the trainees.

**Interpretation and implications:** Trainers and trainees may be reluctant to assign social codes, such as ‘burnout’ or ‘child abuse’, as this might harm their relationship with the patient or even have legal consequences (patients in the Netherlands have the legal right to see their medical notes).

Another explanation for the under-reporting of social codes is that they are perceived to be heterologous ICPC codes and/or ‘second choice’. As a result, social codes are possibly only picked if no other code fits, or in very obvious situations. It thus can be questioned whether the social problems of a population are properly reflected by the ICPC codes of EPR-based data. When asked for situations prone to non-coding on ICPC coding barriers (Chapter 3), social problems were not mentioned spontaneously. Unfortunately, we did not actively ask whether the trainees and trainers experienced coding social problems as a barrier.

**Cardiovascular prevention and diabetes mellitus**

**Main findings:** Differences between trainers and trainees found in previous studies regarding circulatory\(^{12; 25-29; 34}\) and metabolic\(^{12; 25; 26; 28}\) diseases were confirmed in our study. When focusing on cardiovascular prevention, the proportion of patients seen by trainees for cardiovascular prevention was about half that of their trainers. With respect to diabetes mellitus, the proportion of patients seen by first-year trainees was half of that seen by their trainers; for third-year trainees, the proportion of patients seen by the trainers and trainees was comparable.

**Interpretation and implications:** The trainee/trainer differences regarding cardiovascular prevention indicates under-exposure; however, the exposure was still substantial, so competence development may be unabatedly possible. A potentially more problematic situation regards diabetes mellitus. The majority of the training practices included in the study had a nurse practitioner (NP) working in the practice. This is in line with the result, as was found in our national survey among medical receptionists (86% of the practices had an NP; Chapter 6). NPs hold consultations of patients with diabetes mellitus under the final responsibility of the GP, so competence of the GP is required.\(^{35}\) Since the entry of NPs into primary care, the daily workload of GPs with diabetic patients diminished, as did the daily practice of their trainees.\(^{36}\)
Training programmes should respond to the threat of too little diabetes experience, which may result in insufficient competence, and to the consequences of this relatively new type of care organization. Tailored training and interventions directed at the participation of the trainee in the NP’s surgery hours, and at the supervising sessions that the GP holds with the NP should be considered. The experience of GPs and their trainees with diabetes care is not fully reflected by EPR-data extractions, because their NP-supervising activity is not always registered in the EPR, at least not for the trainees. Moreover, as stated before, the software did not count co-morbid conditions; if diabetic patients present with complaints leading to ICPC codes, their diabetic condition may also be considered, but this might not be reflected by the registered code (e.g. dizziness, incontinence, overweight, etc.).

**Acute diseases**

**Main findings:** In Chapter 5, it is shown that trainees in the first year saw fewer acute conditions than their trainers (1.8% vs. 2.3%). The difference was not significant in the third year (2.2% vs. 2.4%). These small differences do not seem to be relevant since the numbers these proportions represent are relatively low.

**Interpretation and implications:** It should be noted that GP trainees acquire most of their experience with acute conditions during their out-of-hours care in central GP cooperatives and during the emergency care rotation in the second year of the training. The EPR extractions that were used in our studies did not comprise those data. Therefore, a complete overview of all acute conditions GP trainees are confronted with during their training could not be provided.

There was no workable cluster of acute conditions available.\textsuperscript{11,12} We therefore composed a cluster of ICPC codes, defined by ‘disease or condition for which diagnostic and/or therapeutic action is required immediately or at most within hours; the seriousness of the condition is not relevant.’ We did not check the exclusive use of these ‘acute’ ICPC codes on the acute moment of presentation, so it is possible that the ICPC code was also used in a non-acute follow-up consultation (e.g. wound control). The low exposure during daytime and the high level of competence that is demanded,\textsuperscript{37} emphasize the importance of the training in the GP cooperatives and of the second-year emergency care rotations. To complete the picture, monitoring of the trainees’ patient mix during their work in the GP cooperatives should be considered.
Minor illnesses

Main findings: The exposure to minor illnesses was almost the same for first- and for third-year trainees (36.3% vs. 35.9%). In line with other studies, trainees saw more minor illnesses than their trainers.

Interpretation and implications: In the first year of training, this does not seem to be a problem, since most trainees have no or very little experience with these illnesses, and this is a year-specific learning goal. By seeing minor illnesses, trainees also gain more experience in their role as a family doctor. Boredom or saturation is unlikely, since the cluster consisted of 278 different ICPC codes (some did overlap) of which the most frequently encountered (R74 Upper respiratory infection acute, R74.1 Common cold and R74.2 Acute pharyngitis) together accounted only for 5.2% of the trainees’ patient mix (see Chapter 5). For third-year trainees, however, this might be a disproportionally large part of the patient encounters, considering the year-specific training goals.

Musculoskeletal, respiratory, and skin conditions

Main findings: Trainees saw musculoskeletal, respiratory, and skin diseases most frequently (Chapters 4 and 5), and the latter two significantly more than their trainers did.

Interpretation and implications: Here, saturation with specific diseases is also unlikely; however, if acute respiratory diseases, acute or chronic sinusitis, acute bronchitis/bronchiolitis, and cough are combined (Table 1, Chapter 5) they roughly account for 10% of the trainees’ patient mix. The learning benefit of this relatively large proportion of acute respiratory conditions can be questioned. Seeing so many respiratory conditions diminishes the possibility of seeing other relevant conditions. Interventions aiming at seeing more low-exposure conditions can eventually be combined with seeing fewer acute respiratory conditions. By doing so, the risk that the trainees’ surgery hours cannot be filled is small, as 65.4% of the receptionists ‘never’ or ‘seldom’ found it hard to fill the trainees’ surgery hours, (unreported data belonging to the same questionnaire as in Chapter 6). In addition, diminishing the proportion of skin and musculoskeletal diseases seems less expedient, since the frequently encountered ICPC diagnoses (dermatomycosis, low back pain, see Table 2, Chapter 4 and Table 1, Chapter 5) are not encountered at the same scale as the acute respiratory conditions.
CHAPTER 6: REASONS FOR DISPARITIES; WHICH ASSIGNING BEHAVIOUR DO RECESSIONISTS REPORT?

Main findings: To find an explanation for the disparities in patient mix between trainers and trainees, we studied the assigning behaviour of the medical receptionist. Almost all receptionists asked patients about their reason for consultation and the majority usually discussed which doctor would be seen. Receptionists were not inclined to assign a patient specifically to either the GP trainer or trainee. Most agreed with the statement that the patient mix of trainees and trainers is similar.

Interpretation and implications: Our findings suggest that most receptionists have at least some freedom when assigning patients. Their assigning strategy was not only passive (assigning any patient to any doctor), but they also reported intentionally assigning a varied patient mix to trainees. The findings in Chapter 6 are in contrast with most disparities in patient mix that we found in Chapters 4 and 5. An explanation for this can be that receptionists have limited insight into the kind of health problems that they assign; the reason for the encounter does not always correspond to the diagnosis that is made. For example, a sinusitis diagnosed in a patient presenting with headache or a depressed patient presenting with vague physical complaints. This explanation is confirmed by the results of the steering intervention (Chapter 7, see further); receptionists reported actively steering substantial numbers of skin conditions and, to a lesser degree, psychosocial conditions. This was not objectified by an increased patient volume in both steered categories. These findings also cast serious doubt on the insight receptionists have or can have on the resulting patient mix they assign. This was not found in all areas, because many receptionists reported to have a preference to assign ‘complex’ patients to the trainer and minor ailments to the trainee (Chapter 6), which is more or less confirmed by our findings in Chapter 5. It thus must be concluded that the receptionist does have some influence on the trainees’ patient mix in a few, but certainly not all, areas. Their insight into the resulting patient mix is, nevertheless, limited.

CHAPTER 7: IS IT POSSIBLE TO INTENTIONALLY STEER TRAINEES’ PATIENT MIX BY INSTRUCTING RECESSIONISTS, TRAINERS, AND TRAINEES?

Main findings: We conducted a RCT to investigate whether a trainee’s patient mix could be actively steered. We carefully concluded that steering of skin conditions was possible. Although, in absolute number, only a non-significant trend could be observed, there was a higher percentual increase in the intervention condition com-
pared with the control condition. Comparatively, the volume of psychosocial-conditions seen did not increase, neither in absolute numbers nor proportionally.

**Interpretation and implications:** The striking contrast between the disappointing steering success and the intervention group receptionists’ perceptions of their own steering is commented on in the previous section. A major limitation was that the intervention had not been tailored to the learning needs of the individual trainee. Such an approach would require the development of multiple learning outcome instruments and this was not considered feasible in this study. Our untailored approach may have hampered the motivation of trainer and trainee for steering.

A second limitation concerned the relatively short duration of each intervention. One trimester was probably too short to see a major effect. Since a non-significant trend towards a higher volume of skin diseases was observed, extending the steering period might have revealed clearer results. It remains questionable, though, how long receptionists could be motivated to continue steering.

**DOES STEERING CONTRIBUTE TO BETTER LEARNING?**

**Main findings:** In Chapter 2, patient mix was found to be positively related to self-reported learning outcome. In Chapter 7, we studied whether this relationship could be demonstrated if the patient mix was actively steered. The learning effect of steering was assessed using self-efficacy and knowledge. However, as described before, active steering was not found to be entirely successful, despite a solid intervention. The higher percentual increase of skin diseases in the intervention group did not result in greater self-efficacy. For psychosocial conditions, a higher self-efficacy was found in the intervention group. Knowledge of psychiatric conditions neither increased between pre- and post-testing, nor between the intervention arms. In a regression analysis, patient volume was a significant predictor of both skin and psychosocial self-efficacy.

**Interpretation and implications:** These findings confirm, at least partly, the results of the systematic review (Chapter 2), that the volume of exposure was related to self-reported outcome measures (self-efficacy) but not to formal assessment (the knowledge test). The knowledge test was not part of the summative assessment, so preparation bias, in contrast to several studies in the review, was unlikely. Perhaps experience increases confidence more than competence,38-40 and the individual idea of being more competent by experience can be deceptive.
3. General educational implications for the GP specialty training

GENERALIZATION OF CLINICAL COMPETENCE AND MINIMUM STANDARDS OF CLINICAL EXPERIENCE

Earlier studies have shown the importance of content-specific knowledge. This may implicate that, ideally, trainees should be exposed to every thinkable health problem to be fully competent. This is practically impossible, so the question is which conditions should at least be encountered, and how often, to acquire competence? Clinical competences may generalize to other competences because they follow similar medical, biological, or psychological patterns. Because of the high diversity of diagnoses encountered in general practice, general diagnostic ability is indispensable. GPs must be able to diagnose, and start the initial treatment of diseases, without prior experience with each distinctive diagnosis.

Minimum requirements may be desirable for trainers, trainees and the training institute because they offer a grip on the quantity that has to be met to expect competence. They can be based on reference data or on consensus by experts in the medical educational/GP training field. However, learning growth curves have been shown to be individually different and exposure never guarantees competence - which must be assessed by other means. Minimum standards may, therefore, be helpful, but the opportunity to see more if needed must be left open, if the required competence is not yet reached.

Since indications of relative ‘over-exposure’ were found, one can also think of formulating maximum standards. The competence growth curve may have reached the flattened part, and additional cases would not contribute further to learning or may even compromise other areas of learning that should be covered. Optimally, the qualification requirements are composed of a portfolio consisting of the encountered patient mix in combination with theoretical and practical assessments.

COMPARING THE PATIENT MIX OF GP TRAINEES WITH TRAINERS OR NATIONAL MORBIDITY FIGURES

As explained in the introduction, relatively old studies and small-scaled studies indicated gaps in the patient mix of GP trainees and disparities between GP trainer and trainees. In Chapters 4 and 5, it is expounded that, based on the results of our large, longterm cohort study, the assumed disparities do indeed ex-
ist. However, an identical patient mix was never explicitly aimed for at our training institute, and probably not elsewhere either. Disparities in the patient mix between trainers and trainees are relevant if they have implications for the competence development of the trainees. The optimal patient mix can be highly different between individual trainees due to earlier working experiences and differences in learning style, learning speed, and learning growth curve.

At the group level, the comparison between the patient mix of the trainees and trainers is informative with respect to the question as to what could be seen in average training practices. It is doubtful, however, whether the trainers’ patient mix is the best reference standard to compare the trainees’ patient mix, when investigating its appropriateness. Since the patients are shared between trainer and trainee, differences between them are often larger, when compared with national figures on the patient mix of the average GP. This last comparison would perhaps be more realistic when discussing under- and over-exposure. An attempt to do this was made in Chapter 5. Conclusions about these comparisons must be made cautiously, because the data, due to differences in extraction method, are not fully comparable.

MONITORING AND STEERING THE PATIENT MIX OF GP TRAINEES
In this thesis, it is shown that EPR-based monitoring provides insight into the patient mix of GP trainers and trainees. Our analyses were performed at the group level, but individual reports could provide specific feedback on many important aspects of the training that have educational value. This includes background numbers, number of telephone or face-to-face consultations and home visits, specific diseases (hypothyroidism, asthma, conjunctivitis) and clusters of disease (malignancies, chronic diseases), diagnostic measures (laboratory, radiology) and therapy (pharmacologic prescriptions, referrals). Continuous monitoring of the patient mix provides a tremendous amount of didactic possibilities. The black box the training practices were once (and still often are) in can be permanently opened. The actual experience of trainees, which can be seen as part of their portfolio, becomes available for evaluation.

Providing patient mix reports to trainers and trainees, will confront them with disparities from the earlier mentioned reference standards, based on averages or on consensus. If confronted with these disparities, trainers and trainees will start questioning whether their diagnoses are correct, whether they use the right ICPC
chapter 8

codes, whether the trainee sees enough patients, whether the training practice suits their learning needs, whether they should take part in the surgery hours of the NP more often, among others. This information thus increases awareness and will most probably lead to a variety of actions intended to tailor the patient mix to the learning needs of the trainee, aiming for an optimal amount of experience in the specific domains needed for their training.

Achieving this by active steering of the patient mix was only moderately successful as was shown in Chapter 7. Instead of steering as many patients as possible, as was the instruction of our interventions, future steering should be considered to be directed at target values which are then permanently monitored. Both trainer and trainee should be responsible for reaching the targets, the receptionist can be involved. In this way, cooperation is encouraged. The aim of the intervention then is clearer and the participants are probably more motivated for steering.

Considering the role of the receptionist, the finding that trainees see younger patients is relevant. If, for instance, it is desired that more chronic conditions are seen, it would be much easier to instruct receptionists to assign, to the trainee, patients of an age in which the chance of encountering the interventional condition (in this case, chronic conditions) is higher (e.g. 70+ years), than instructing the receptionist to assign more chronic conditions. This could also be implemented in training practices where the reason for the consultation is not commonly asked by the receptionist.

When a monitoring system is implemented, it will be interesting to study which of the various actions that will take place to balance a patient mix are most effective. The effect of the EPR-based data reports may be enhanced if (minimum) standards are required for finishing the training. The data could then be part of the trainee’s assessment and could be linked to examinations based on, for instance, entrustable professional activities (EPAs).

Some conditions, such as diseases of the blood and blood-forming organs present infrequently, also to regular Dutch GPs. From an educational perspective, potential under-exposure in these areas is conceivable. In Chapter 2, it is argued that supervision is important for competence development. Hence, if it is not possible to increase the number of patients in these areas, special supervised sessions for low-exposure conditions should be considered; however, the effectiveness of this approach should be studied.
Training institutes can benefit from aggregated EPR-extracted information. The curriculum can be adjusted to the patient mix – and the gaps this contains. In addition, the teaching programme can be better adjusted to the actual patient mix. Trainees should build expert performance in the areas they will frequently encounter. For low-exposure areas, learning aims should be realistic, involving a minimum-competence package, concentrating on initial treatment and on not missing important diagnoses.

VALIDITY OF ICPC BASED PATIENT MIX DESCRIPTIONS

As explained in Chapter 3, patient mix descriptions based on EPR-derived data are valid for research and educational purposes.\(^{55-57}\) Many earlier studies use handwritten or digital logbooks that are intended solely for educational purposes\(^{47;58;59}\) and thereby subject to social-desirability bias. Extracted from EPR systems, our data were not biased by any threat of formal assessment and did not acquire actions apart from the usual patient care. We therefore believe that EPR-derived logbooks are superior to traditional logbooks that are kept separately.

The validity of the EPR-based derived logbooks is, however, compromised by the ICPC classification system they are built on, which was found to be insufficiently refined\(^{60}\) (Chapter 6). Not all GPs used the third ICPC digit which provides some more options for detailed classification. Furthermore, ICPC codes themselves do not reflect the stage of the disease or its severity. By agreement, ICPC codes are used at ‘the highest level of specificity’.\(^{61}\) This means that the choices for codes are, or should be, made conservatively, and, thus, information may be lost. The validity is also dependent on the diagnostic competence of the trainees and the adequacy of the attribution of the right codes to the diagnosis given.\(^{8}\) As already mentioned, not all co-morbidity and differential diagnostic considerations are reflected by the ICPC diagnosis.

Another major drawback is that the ICPC classification system was not developed for educational purposes. To obtain a better perception of the patient mix and to focus on the educational value, clusters of diseases can be useful. The only existing cluster we found appropriate for our descriptions was chronic diseases, composed by Knottnerus et al.\(^{62}\) More clusters were published but not appropriate for our data for different reasons (based on ICPC-2,\(^{63}\) no clear educational meaning\(^{64}\)). We experienced that the composition of meaningful educational clusters of disease was particularly challenging (Chapter 5).
Although all trainees were affiliated with the same training institute, we do not think that this influenced the results, as other studies in different times, scales, and areas have shown similar results.

The use of data from EPR systems combined with decision rules to aggregate the data logically and consistently can be of great educational value. However, some of these decision rules were based on pragmatic grounds. For example, if diagnosis codes were linked to episodes of care, the episode code was counted only if no other new diagnosis code was entered during that contact. The episode code was ignored if a newly entered ‘distinct’ diagnosis code was entered and linked to the episode code. This choice was made because, in the pilot study, it was found that the use of episode codes highly differed between participating trainers. If, in future studies, both the episode and the newly entered ‘distinct’ codes are counted separately, it could contribute to a better understanding of co-morbidity and consultation behaviour of, for instance, patients with complex or chronic diseases. Other choices we made concerning the decision rules were equally defendable, but also based on pragmatic grounds.

4. Recommendations for future research
Future studies should explore the possibility of minimum, maximum, and optimal standards for the exposure of several domains during GP specialty training. For this kind of research, a permanent monitoring system in combination with repeated competence assessments is desirable.

A further question is whether tailoring patient mix of individual GP trainees to their learning goals is effective for learning, and how this can be best effectuated. This way, the relationship between patient mix and learning outcome can be mapped out further, and the specific contribution of several aspects of patient mix to learning can be identified.

EPR-based monitoring has more meaning if the patient mix descriptions are based on meaningful educational clusters. There is, to date, no consensus on the composition of many of these clusters. Developing them is a challenge and it would, indeed, be interesting to study which of these clusters have educational meaning.

5. Final word
For decades it was presumed that the patient mix of GP trainees contained low-exposure areas which had potentially negative consequences for the trainees’ com-
petence development. This thesis provides insight into the black box in which the patient mix had been. The low-exposure areas have been identified. The implementation of a patient mix monitoring instrument offers the possibility of tailoring the patient mix to the individual needs of GP trainees and the opportunity to answer many educationally relevant research questions.
REFERENCES

2. Health Care Supply General Practice 2009 [Aanbod Huisartsenpraktijken 2009]. Dutch College of General Practitioners (NHG) and Dutch Association of General Practitioners (LHV).
4. Statistics Netherlands [Centraal Bureau voor de Statistiek]. www.cbs.nl
17. Pham HH, Simonson L, Elnicki DM, Fried LP, Goroll AH, Bass EB. Training U.S.
Unlimited Exposure


37. NHG position paper. The contribution of the GP in the chain of care for acute diseases


57. Lyman JA, Schorling J, Nadkarni M, May N, Scully K, Voss J. Development of a Web-based resident profiling tool to sup-


64. Dutch Institute for Health Services Research (NIVEL) http://www.nivel.nl
CHAPTER
9
Summary
Summary

Chapter 1
Earlier research suggests that the patient mix of GP trainees differs from their trainers and contains areas of low exposure, which may have consequences for the trainees’ competence development. This thesis focuses on the composition of the patient mix of 73 GP trainees and their trainers. In Chapter 1 the specific characteristics of the patient mix of Dutch GPs and the setting of the GP training are portrayed. Subsequently, the term ‘patient mix’ is defined and an important learning theory, deliberate practice, is presented. In Chapter 1 we introduce our method to get a clear picture of the patient mix of GP trainees. We used data extracted from electronic patient records (EPRs) to identify areas of low exposure and disparities between GP trainees and their trainers. Next we propose the possibility of increasing the exposure by actively steering the trainees’ patient mix.

Chapter 1 concludes by stating the research questions and by giving an exposition of the methodology of the CASANOVA (CASE AssigNment fOr GP VocTional training) project.

Chapter 2
In a systemic review, the empirical evidence about the contribution of patient mix in work-based learning is addressed. We conducted a literature search across Medline, Embase, Web of Science, ERIC and the Cochrane Library. Original, quantitative studies were included. After screening of 10,420 citations, 298 full-text articles were retrieved of which 22 were identified as relevant. Two reviewers extracted data using a coding sheet, and assessed the methodological quality of the studies. Operationalizations of patient mix were fairly different, allowing very few inferences from the studies. Learning outcomes were divided into self-reported and formal assessment.

We found a positive relationship between patient mix and self-reported outcomes evaluating the progress in competence as experienced by the trainee, such as self-confidence and comfort level. Patient mix was also found to correlate positively with self-reported outcomes evaluating the quality of the learning period, such as self-reported learning benefit, experienced effectiveness of the rotation, or the instructional quality. Supervision and learning style seemed to be mediating variables of the relationship between patient mix and learning.

A relationship between patient mix and formal assessment was rarely dem-
onstrated. Presumably, increasing experience only, does not automatically lead to more competence. However, most formal assessments were not study-specific and therefore probably not sensitive enough to ascertain how patient mix can affect learning outcome.

Chapter 3
To gain insight into the validity of our patient mix descriptions, the coding behaviour of the trainers and trainees was investigated in a questionnaire. Motivation-related, ICPC-system-related and external barriers of GP trainees and trainers while using ICPC coding were asked for. Results were compared with self-reported coding performance and actually measured EPR-based coding performance. Hence, coding bias, due to the barriers that GP trainers and trainees experience while using the ICPC classification system, was addressed.

The percentages of contacts with an ICPC code were high, both for GP trainers and for trainees. Trainers’ estimation of their coding performance was rather accurate, considering the high correlation with the actual coding percentage. For trainees this correlation was lower. Most participants reported always to register ICPC codes for consultations and home visits. Telephone consultations, repeat prescriptions and administrative actions were coded less frequently. Most participants never or rarely experienced coding barriers, an exception being the ‘insufficient refinement’ of the ICPC system. Most motivation and ICPC-related barriers correlated with self-reported and actual coding performance. In regression analyses ‘ICPC coding is unpleasant to use’ was found to predict the trainees’ coding percentage. This also predicted the trainers’ coding percentage, as did ‘no personal gain from ICPC’ and ‘coding is difficult’ (total variance explained was 41%). As no other relevant areas were prone to non-coding, EPR-derived data seem biased by non-coded telephone consultations only.

Chapter 4
To describe the patient mix of GP trainees and to study differences in patient mix between first-year and third-year GP trainees, a cohort study was started in 73 GP training practices. During a six-month period, aggregated EPR data about ICPC diagnosis codes and data on the gender and age of all patient contacts were collected. We extracted data directly from the EPR every three months. The raw extraction data were processed by multiple software routines. We programmed decision rules to ensure logic and consistency in counting the codes.
The number of patient contacts was not significantly different between the first- and third year trainees, but the standard deviation in the third year doubled. This was largely due to several trainees who saw exceptionally large numbers of patients.

Third-year trainees saw more patients between 45 and 64 years than did first-year trainees. Further differences were found between years 1 and 3 for diseases of the blood and blood-forming organs, psychiatric diseases and metabolic diseases. The mean number of chronic patients seen per trimester was 48.0 (SD 21.9) for first-year trainees and 62.4 (SD 32.5) for third-year trainees. Female trainees saw twice as many female conditions as male trainees did per trimester. Differences in patient mix between first- and third-year trainees seem at least partly related to year-specific learning objectives. The use of an EPR derived educational instrument provides a useful insight into the trainee’s patient mix.

Chapter 5
The results of the EPR based extractions of the 73 trainees were compared with the aggregated data of their 114 GP trainers. National morbidity figures of LINH 2009 and the second Dutch National Survey of General Practice were used as an external reference.

The results showed that trainers had double the number of face-to-face consultations, and treble the number of telephone consultations compared to trainees. Trainees’ patient mix consisted of significantly more patients with eye diseases, ear diseases, respiratory diseases, skin diseases and minor illnesses compared to their trainers. Trainers encountered significantly more patients with circulatory diseases, psychiatric diseases, metabolic diseases, male genital conditions, social conditions, chronic and oncological diseases. Female trainers and trainees encountered almost twice as much female conditions compared to their male compeers. Male conditions were seen more often by male doctors.

In conclusion, trainers and teachers at the specialty training must be aware of areas of low exposure. Trainers should ensure trainees handle more chronic, complex, psychosocial and circulatory conditions.

Chapter 6
In search for an explanation for the disparities between GP trainees and trainers found in chapter 5, the attitude of patients towards trainees has been extensively
studied. The assigning behaviour of the medical receptionist is another important factor that may be accountable for the disparities. We identified topics in the assigning behaviour of receptionists in a focus group. Subsequently, we designed a questionnaire to study assigning behaviour. The questionnaire was sent to 478 GP training practices in the Netherlands, 326 were returned (68%).

Most receptionists routinely asked for the reason for the consultation. The majority at least sometimes discussed with patients which doctor would be seen. Most receptionists have at least some latitude when assigning patients to either the trainer or the trainee. Most receptionists considered the patient mix of trainees and trainers to be similar. Almost all receptionists reported ‘often’ or ‘always’ assigning ‘every possible problem’ to the trainee. A similar picture arose regarding specific subpopulations. However, the receptionists reported that they assigned complex and new patients to the trainers more often than to the trainees. In conclusion, with some exceptions, receptionists try to assign trainees a varied patient mix.

Chapter 7

This chapter investigates whether patient mix can be steered by instructing medical receptionists, trainers and trainees in a randomized controlled trial. The 73 trainees were randomized into an intervention group and a control group. During two successive periods of three months, patients with skin conditions and psychosocial conditions were actively assigned to trainees in the intervention group (n=37). The patient mix was measured by EPR extracted data and compared with a baseline registration (Chapter 4 and 5). Learning outcome was measured by self-efficacy questionnaires and by a knowledge test.

No difference in increase was found in patient volume and diversity of the steered conditions in the intervention group as compared to the control group. However, the percentual increase of exposure to skin conditions was greater in the intervention group. No difference in skin conditions self-efficacy and psychiatric knowledge was found. The increase of psychosocial self-efficacy however, was greater in the intervention group. In regression analyses, patient volume was found to be a significant predictor of both skin and psychosocial self-efficacy. We conclude that steering the patient mix of a GP trainee is not as easy as it seems. We suggest to study the effects of tailoring patient mix to individual learning aims and to investigate the effects of predefined prerequisite exposure numbers on steering.
Chapter 8

In the first part of the general discussion the main findings are summarised for each chapter, followed by an interpretation of the results and the implications for the GP specialty training. In the second part of this chapter general educational implications for GP specialty training are discussed. Individual feedback reports, if combined with reference standards, either based on consensus or on group averages, will confront trainers and trainees with disparities. We expect that this approach leads to a variety of actions intended to tailor the patient mix to these reference standards, and to the individual learning needs of the trainee if the data are also linked to the trainee’s assessments. Whether this is effective for learning indeed should be further studied. For this, the implementation of a patient mix monitoring system is essential.
CHAPTER 10

Samenvatting
Samenvatting

**Hoofdstuk 1**
Uit eerder onderzoek is gebleken dat er mogelijk verschillen bestaan tussen de gezondheidsproblemen die artsen in opleiding tot specialist huisartsgeneeskunde (aios) en huisartsopleiders op hun spreekuur te zien krijgen. Mogelijk zijn er gebieden in de ‘patiëntenmix’ van aios die gedurende hun opleiding onderbelicht blijven, en dat kan consequenties hebben voor hun competentieontwikkeling. Dit proefschrift richt zich op de samenstelling van de patiëntenmix van 73 aios huisarts-geneeskunde en hun opleiders.

In hoofdstuk 1 worden de specifieke kenmerken van de patiëntenmix bij huisartsen in Nederland geschetst. Vervolgens wordt de term patiëntenmix gedefinieerd en wordt een belangrijke onderwijskundige theorie, ‘deliberate practice’, geïntroduceerd. Daarna wordt de methode gepresenteerd waarmee we de patiëntenmix in huisarts(opleidings)praktijken in beeld hebben gebracht. Hiervoor is gebruik gemaakt van gegevensextracties uit het huisartseninformatiesysteem (HIS). Zo doende konden we gebieden identificeren in de patiëntenmix die in de opleiding mogelijk onderbelicht blijven. Vervolgens bespreken we de mogelijkheid om het patiëntenaanbod op deze gebieden door middel van actieve sturing te vergroten. Het hoofdstuk eindigt met een opsomming van de onderzoeksvragen en een uiteenzetting van de methodologie van het CASANOVA (CASe AssigNment fOr GP VocAtional training) project.

**Hoofdstuk 2**

Er bleken grote verschillen tussen studies te bestaan in de manier waarop patiëntenmix werd geoperationaliseerd, waardoor de mogelijkheden tot gevolg-
trekkingen beperkt bleven. De in de studies gerapporteerde leeruitkomsten waren ofwel gebaseerd op zelfevaluatie, ofwel op formele toetsen, die meestal onderdeel uitmaakten van het standaard toetsprogramma van de betreffende opleiding.

We vonden een positief verband tussen patientenmix en leeruitkomsten die de zelf-ervaren vooruitgang in competentie weerspiegelen, zoals de mate van zelfvertrouwen. We vonden ook een positieve relatie tussen patientenmix en zelfgerapporteerde uitkomsten die de kwaliteit van de opleidingsperiode beoordeelden, zoals leeropbrengst, ervaren effectiviteit of de instructieve kwaliteit. De kwaliteit van de supervisie en de leerstijl van de student bleken van invloed op deze relaties.

Een relatie tussen de patiëntmix en de prestaties op formele toetsen werd nimmer duidelijk aangetoond en dus is het aannemelijk te veronderstellen dat meer ervaring niet automatisch tot meer competentie leidt. De meeste formele toetsen in de artikelen waren echter niet speciaal ontwikkeld voor de betreffende studie en daarom mogelijk niet sensitiief genoeg om te ontdekken hoe de patiëntmix het leren beïnvloedt.

Hoofdstuk 3
Om een beter beeld te krijgen van de validiteit van onze operationalisaties van patiëntmix onderzochten we het codeergedrag van aios en opleiders met behulp van een vragenlijst. Aios en opleiders werd gevraagd in hoeverre hun codeergedrag werd beïnvloed door motivatie-gerelateerde barrières, aan het ICPC-systeem gerelateerde barrières, en door barrières die voortkomen uit externe factoren. De resultaten vergeleken we met de zelfgerapporteerde codeerprestaties van de respondenten en met de daadwerkelijke codeerprestatie, gemeten met behulp van HIS extracties. De vertekening in de patiëntmixbeschrijvingen, die zou kunnen ontstaan als gevolg van de barrières die aios en opleiders ervaren als ze met het ICPC-systeem coderen werd zodoende inzichtelijk gemaakt.

De codeerpercentages waren hoog, zowel onder aios als bij de opleiders. Opleiders bleken een redelijk goede inschatting te maken van hun codeerpercentage, gezien de hoge correlatie tussen het zelfgerapporteerde en het werkelijk gemeten percentage. Bij aios was deze relatie minder sterk. Telefonische consulten, herhaalrecepten en administratieve handelingen werden minder frequent gecodeerd. De meeste aios en opleiders ervaarden geen of weinig barrières, behalve met de grofmazigheid van het ICPC systeem. De meeste motivatie- en ICPC-gerelateerde barrières correleerden in hoge mate met de zelfgerapporteerde en de daadwerkelijk geme-
ten codeerpercentages. Uit regressieanalyses bleek de barrière ‘Ik vind het onprettig om met de ICPC te moeten werken’ het werkelijke codeerpercentage van aios te voorspellen. Deze barrière voorspelde ook het gemeten codeerpercentage van de opleiders, evenals de barrières ‘als het beter betaald zou worden, zou ik meer code- ren’, en ‘coderen vind ik moeilijk’, (de totale verklaarde variantie van deze regressie was 41%). Omdat het wel of niet coderen van herhaalrecepten en administratieve handelingen weinig relevant is voor de beschrijving van de patiëntenmix, concluderen wij dat op HIS gebaseerde patiëntenmix-beschrijvingen alleen vertekend worden door ongecodeerde telefonische contacten.

**Hoofdstuk 4**

Hoofdstuk 4 beschrijft de patiëntenmix van eerstejaars en derdejaars aios en verschillen tussen deze opleidingsjaren worden gerapporteerd. Hiertoe deden we een cohortstudie in 73 opleidingspraktijken. Gedurende 6 maanden verzamelden we gegevens over ICPC-diagnosecodes en leeftijd en geslacht van patiënten. Iedere 3 maanden werden data uit de HISsen geëxtraheerd. Met behulp van een softwareapplicatie werden de data op consistentie en logische wijze geaggregeerd.

Het aantal patiëntencontacten verschilde niet significant tussen eerste en derdejaars aios, maar de standaarddeviatie (SD) van het aantal consulten in het 3e jaar was twee keer zo groot als in het eerste jaar. Dit bleek te worden veroorzaakt door enkele aios die uitzonderlijk veel patiënten zagen.

Derdejaars aios zagen meer patiënten tussen de 45 en de 64 jaar dan eerstejaars. Verder verschilde het aantal patiënten met ziekten van het bloed of bloedvormende organen, psychiatrische aandoeningen en metabole ziekten. Het gemiddeld aantal chronische patiënten was 48.0 (SD 21.9) in het eerste jaar en 62.4 (SD 32.5) in het derde jaar. Vrouwelijke aios zagen per trimester gemiddeld twee keer zoveel gynaecologische en/of zwangerschapsgerelateerde aandoeningen als mannelijke aios. De verschillen tussen het eerste en derde jaar kunnen deels berusten op jaar-specifieke leerdoelen. Tot slot concluderen we dat een op HIS-extracties gebaseerd onderwijskundig instrument inderdaad een goed beeld geeft van de patiëntenmix van aios.

**Hoofdstuk 5**

In dit hoofdstuk worden de geaggregeerde data van HIS-extracties van 73 aios vergeleken met de data van hun 114 opleiders. Nationale morbiditeitscijfers van LINH 2009 en de 2e Nationale Studie gebruikten we als externe referentie.

Opleiders en docenten van de huisartsopleiding dienen op de hoogte te zijn van de gebieden waarop het aanbod voor aios relatief laag is. Opleiders kunnen worden aangemoedigd om ervoor te zorgen dat hun aios meer chronische, complexe, psychosociale en cardiovasculaire aandoeningen zien.

**Hoofdstuk 6**

De rol van de patiënt als mogelijke oorzaak van de verschillen in patiëntenmix tussen aios en opleiders is in eerdere studies uitvoerig bestudeerd. Het toewijzingsgedrag van de praktijkassistent(e) is een andere factor die mogelijk bepalend is voor deze verschillen. Wij onderzochten het toewijzingsgedrag van praktijkassistentes met een vragenlijst waarvoor we eerst thema’s vaststelden in een focusgroep met praktijkassistentes. De resulterende vragenlijst werd verzonden naar 478 opleidingspraktijken in Nederland, waarvan er 326 (68%) ingevuld werden teruggestuurd.

Uit de antwoorden bleek dat het grootste deel van de assistentes routinematig naar de contactreden van de patiënt vraagt. De meerderheid besprak met de patiënt, variërend van soms tot altijd, bij welke dokter deze wilde komen. Dat betekent dat praktijkassistentes inderdaad de verdeling van patiënten tussen de aios en de opleider kunnen beïnvloeden. De praktijkassistentes plaatsten alle soorten problematiek bij de aios en de meesten vonden de patiëntenmix van aios en opleiders vergelijkbaar. Ook wat betreft specifieke patiëntencategorieën zoals ouderen, kinderen, passanten en spoedgevallen bleken ze nauwelijks een voorkeur te hebben voor routinematige toewijzing aan ofwel de aios of aan de opleider. Complexe en nieuwe patiënten werden echter wel wat vaker bij de opleider geplaatst. Behoudens deze uitzonderingen kan worden geconcludeerd dat praktijkassistentes een zo divers mogelijke patiëntenmix voor de aios proberen te creëren.
**Hoofdstuk 7**

In dit hoofdstuk wordt onderzocht of de patiëntenmix van de aios bijgestuurd kan worden door een instructie te geven aan de praktijkassistenten, de aios en de opleider, in een gerandomiseerde, gecontroleerde studie. De 73 aios werden gerandomiseerd in een interventie en een controlegroep. In de interventiegroep (n=37) werden gedurende 3 maanden patiënten met huidaandoeningen actief op het spreekuur van de aios geplaatst. Dit werd gevolgd door een tweede periode van actieve sturing op psychosociale problematiek. Het aantal patiënten met huidaandoeningen en psychosociale problematiek dat de aios zagen werd gemeten met behulp van HIS-extracties en vergeleken met een baseline meting (zie hoofdstuk 4 en 5). Uitkomstmaten waren huid- en psychosociale self-efficacy, gemeten met een vragenlijst, en psychiatriekennis, gemeten met een online toets.

Tussen de interventiegroep en de controlegroep vonden we geen verschil in toename van patiëntenvolume en diagnose-diversiteit van de gestuurde condities. De percentuele toename van huidziektes was echter wel groter in de interventiegroep.

We vonden geen verschil in toename van huid self-efficacy en psychiatriekennis tussen de onderzoeksarmen. De toename van psychosociale self-efficacy was echter groter in de interventiegroep. In een regressieanalyse bleek het patiëntenvolume een significante voorspeller voor zowel huid- als psychosociale self-efficacy. We concluderen dat het sturen van de patiëntenmix van aios nog niet zo eenvoudig is als het misschien lijkt. Een suggestie is om de patiëntenmix aan te passen aan de individuele leerdoelen en de effecten te onderzoeken van sturing op basis van vooraf gestelde minimum patiëntenaantallen.

**Hoofdstuk 8**

In het eerste deel van de discussie worden de belangrijkste bevindingen per hoofdstuk gepresenteerd, gevolgd door een interpretatie en de implicaties voor de huisartsopleiding. In het tweede deel worden meer algemene onderwijskundige implicaties voor de huisartsopleiding besproken. In dit deel wordt uiteengezet hoe afwijkingen van referentiewaarden (gebaseerd op consensus of op gemiddelden) inzichtelijk worden door individuele feedback-rapportages. Het is te verwachten dat aios en opleiders hierop zullen reageren met een diversiteit aan handelingen, bedoeld om de patiëntenmix aan deze referentiewaarden te laten voldoen. Indien de feedbackrapportage daarnaast wordt gekoppeld aan de prestaties op gerelateerde
toetsen, dan vormt dit een krachtig instrument om de individuele leerbehoeften van de aios inzichtelijk mee te maken. Of dit daadwerkelijk bijdraagt aan effectiever leren moet verder worden onderzocht. Hiertoe is de implementatie van een aios-volgsysteem essentieel.
Dankwoord
“Wolves don’t hunt singly, but always in pairs. The lone wolf was a myth.”

— John Fowles, The Magus
Promotieonderzoek. Het was niet iets waarvan ik had vermoed dat het op mijn pad zou komen. Daar bleken nogal wat mensen; familie, vrienden, collega’s en leidinggevenden anders over te denken -en zie! De grote beslissing om het toch aan te gaan stond –helaas – vooral in het teken van die andere grote beslissing: voorlopig te stoppen als praktiserend huisarts. Daarnaast was ik erg sceptisch over de kans van slagen van het project. Huisartsen aan het ICPC coderen krijgen? Data extraheren uit verschillende, en volstrekt verschillend geïmplementeerde HIS-sen? En dan ook nog een interventiestudie waarbij ook nog eens leeruitkomsten moesten worden gemeten? Het leek me onmogelijk.

Gelukkig waren er mensen die er in geloofden. Diezelfde mensen meenden dat ze het project mij konden toevertrouwen. Welnu, het is gelukt en daar ben ik enorm trots op. Maar het was zeer zeker niet gelukt zonder:

...eerst en vooral mijn promotor, Margreet Wieringa-de Waard. Al vanaf het eerste moment dat we elkaar spraken opperde je de mogelijkheid om promotieonderzoek te gaan doen, ver voordat ik dat zelf serieus heb overwogen. Jij hebt me de kans gegeven, je hebt me altijd het gevoel gegeven dat je het me toevertrouwde en je hebt me altijd alle steun gegeven die ik vroeg of vergat te vragen. Jouw manier van werken heb ik altijd bewonderd: keihard werken, en toch een deur die letterlijk altijd openstaat. Elk stukje tekst dat ik je stuurde becommentarieerde je vrijwel per omgaande. Daarnaast weet je op de een of andere manier altijd de juiste mensen om je heen te verzamelen. Jouw ‘down to earth’ benadering in combinatie met je extreme accuratesse waren leerzaam en waardevol. Hoe je in no time onze onderzoeksgroep hebt opgebouwd is bewonderenswaardig.


Hartelijk dank voor het beoordelen van mijn manuscript en het zitting nemen in de oppositie. Ik verheug me op een aangename en vruchtbare gedachtewisseling.

…alle huisartsopleiders die bereid waren mee te werken, maar vooral alle aios die hebben meegedaan. En bovenal de praktijkassistentes! Laat voor eenieder duidelijk zijn dat jullie een hoofdrol hebben gespeeld in dit onderzoek. Hopelijk zijn jullie ook degenen die er van zullen profiteren.

…Jacob Mohrs: je bent van onschatbare waarde geweest voor de dataverzameling en bewerking. Je stond altijd voor me klaar, en meer dan dat: je wilde voortdurend meedenken. Het was een groot plezier om je zo dichtbij (1 deur verder) en zo laag-drempelig ‘beschikbaar’ te hebben.

…Marga Vintges: ik heb je nooit ontmoet, maar jouw artikel ‘Welke gezondheidproblemen zien huisartsen in opleiding?’ uit 2001 is de basis van dit proefschrift.

…Ernst van der Wiele, Henk van Weert en Henk Brouwer, als ‘founding fathers’ van het hele project: ik ben jullie zeer erkentelijk. Daarnaast wil ik jullie, en natuurlijk ook Bert Schadé, Gerben ter Riet en Patrick Bindels bedanken voor alle steun en tips die ik soms in de wandelgangen, tijdens praatjes, of besprekingen kreeg. Het mag soms zo subtiel lijken, jullie zijn van groot belang geweest.

…ook de andere leden van de begeleidingscommissie: Rien de Vos en Dick Willems dank voor alle wijsheid, intellectuele input, gevraagde en ongevraagde adviezen. Nynke van Dijk: je trad later toe en hebt de gave het te laten lijken of onderzoek doen makkelijk is. Als je jouw adviezen uitvoert komt het goed. Daar zullen nog velen van profiteren!

…Cees van der Vleuten: onze contacten waren kort maar uiterst nuttig. Onze review was hersengymnastiek, ik heb enorm kunnen profiteren van jouw adviezen.

…Machteld IJff: je hebt je keihard ingezet voor de – soms weerbarstige – dataverzameling. Van jou nooit een onvertogen woord als je je door tegenslagen heen moest bewegen. Mijn dank is groot.
…de mensen op de achtergrond: dankzij de kennis en hulp van Thijs Backus werden de eerste HIS-extracties mogelijk. Door jou konden we van start!
…Els Grim: jouw hulp met de data-extracties in Almere was van grote waarde.
…Faridi van Etten-Jamaludin voor de hulp bij het samenstellen van de zoekstrategie en Rudy Scholte voor het maken van de software applicatie.
…Rob Kreuger: dank voor het ontwerp van het logo, het heeft in belangrijke mate bijgedragen aan onze herkenbaarheid – en daarmee denk ik aan het succes van het onderzoek.
…Herman Düsman en Marieke Bednarczyk voor het maken van de psychiatrietoets en de online toets.

…mijn kamergenoten (eigenlijk vooral -tes): Gré, Erik, Birgit, Hélène, Helen, Corlien, Sabine, Jacqueline, Anouk, Marlies en Antje. Ik bof enorm met jullie, stuk voor stuk slim en gezellig. Nooit te beroerd om een tekst even te lezen, mee te denken of een praatje te oefenen. Werken was en is een feest dankzij jullie. Natuurlijk dank ik ook alle andere jonge onderzoekers (JOPpers) en iedereen die zich daaromheen beweegt.

…onze onderzoeksgroep: Sandra, Ria, Ellen en anderen. Heel veel dank voor al jullie adviezen, ik vind het erg waardevol om als onderzoeker niet alleen te staan in het onderzoeks veld. Ik ben er trots op dat we in betrekkelijk korte tijd zo’n professionele club hebben kunnen opbouwen. Dat is een compliment aan ieder persoonlijk. In het bijzonder wil ik Judy bedanken; tegelijk begonnen, met dezelfde begeleiders. Samen in de aula – en dan ook samen het feest.

…alle andere collega´s van de afdeling Huisartsgeneeskunde voor alle hulp en ondersteuning die ik van jullie kreeg.

...iedereen die ik ongetwijfeld vergeten ben.


…Mirjam en Zack: ik ben jullie dankbaar voor het verzinnen van de originele en toepasselijke titel. Super!

…Bas Jacobs: heel erg bedankt voor alle creatieve tijd en wijsheid die je in de layout en het omslagontwerp hebt gestoken. Er ging een wereld voor me open. Ik vind het heel bijzonder dat je dat hebt willen doen en ben erg trots op het resultaat!

…Ben, Anne, Michael, Sarah en Hannah: een geweldige schoonfamilie! Jullie zijn altijd geïnteresseerd in wat wij doen, en kunnen – met ons – enorm genieten van de kinderen. Daar gaan we wat mij betreft mee door.

…Jasmijn: vier handen op een buik. Onze broer-zus loyaliteit is uniek. Ik ben blij met alles dat we, uiteraard samen met Sonja, Saar en Ben doen. Fijn dat je paranimf wilt zijn, heb je het proefschrift al gelezen?

…Sonja en Harry, Will en Anita: jullie hebben me altijd vrij gelaten, bij alles wat ik wilde. Jullie hebben me de ruimte gegeven mijn eigen weg te zoeken, maar jullie zijn er ook altijd als ik daarbij een gesprekspartner zoek. Ik hoop en reken erop dat jullie trots op me zijn.
Lieve Rachel, we zijn al 12 jaar samen. Ik weet dat je me niet gelooft, maar ik het meen het als ik zeg dat ik bijna dagelijks kan genieten van je charme, je oorspronkelijkheid, je humor, je goedheid en je volledige overgave aan het moederschap. Jij hebt van het begin het volste vertrouwen gehad in dat ik dit proefschrift zou kunnen voltooien, en op de momenten dat ik het vroeg gaf je me altijd alle ruimte.

Over de auteur