Embedding trials in evidence-based clinical practice

Oude Rengerink, K.

Citation for published version (APA):
About the author

Katrien Oude Rengerink was born in Oldenzaal in 1984, daughter from Jan Oude Rengerink and Betsie Eppink. Together with her sister and two younger brothers she grew up in Denekamp.

In 2002 she completed atheneum at Twents Carmel College and started her studies Biomedical Sciences at the Radboud University Nijmegen. During her Bachelor training period at the Municipal Health Services (GGD) Den Bosch, she studied smoking behavior of parents of young children. For her management profile training period she reviewed the physiological changes during hunger strike and hunger strike protocols in the Netherlands, at the Johannes Wier Stichting. As part of her minor in International Health she participated in the Community Health Rotation at the Muhimbili University of Health Sciences, Tanzania. She finished her master in epidemiology with the development of a prediction model for recurrent upper respiratory tract infections, at the Julius Center for Health Sciences and Primary Care, supervised by dr. Maroeska Rovers.

After her graduation in 2007 she volunteered in the children’s department of hospital San Juan de Dios in Arequipa, and travelled Peru and Bolivia. In 2008 she started her PhD at the Department of Obstetrics and Gynaecology of the Academic Medical Center, University of Amsterdam, under supervision of Professor Ben Willem Mol, Professor Patrick Bossuyt and doctor Lotty Hooft. She lives with René in Utrecht.
Embedding trials in evidence-based clinical practice

Katrien Oude Rengerink
Embedding trials in evidence-based clinical practice
PhD thesis, University of Amsterdam, the Netherlands, 2014

The work presented in this thesis was performed at the Department of Obstetrics and Gynaecology of the Academic Medical Center (AMC), University of Amsterdam.

The IMPACT project has been funded by ZonMw Grant no. 80-82310-97-10210. The EU-EBM project has been funded by the European Union Leonardo da Vinci Vocational Training Action Programme, Grant no. UK/05/B/F/PP-162_349

Financial support for printing this thesis was provided by Roche Nederland B.V. Woerden

Copyright ©Katrien Oude Rengerink, Amsterdam, the Netherlands, 2014. All rights reserved. No part of this thesis may be reproduced, stored or transmitted in any form or by any means, without prior permission of the author.

Cover Janne Oude Rengerink
Lay-out Janne Oude Rengerink
Printed by Ridderprint
Embedding trials in evidence-based clinical practice

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 Agnietenkapel
op woensdag 21 mei 2014, te 14.00 uur

door

Katrien Oude Rengerink

geboren te Oldenzaal
Promotiecommissie

Promotores
Prof. dr. P.M.M. Bossuyt
Prof. dr. B.W.J Mol

Co-promotor
Dr. L. Hooft

Overige leden
Prof. dr. F. van der Veen
Prof. dr. M.A. Boermeester
Prof. dr. G.P. Westert
Prof. dr. R.J. de Haan
Prof. dr. R.J.P.M. Scholten
Prof. dr. T.J. ten Cate

Faculteit der Geneeskunde
## Contents

### Chapter 1 General introduction 8

### PART 1: IMPROVING RECRUITMENT IN CLINICAL TRIALS

### Chapter 2 IMproving PArticipation of patients in Clinical Trials – rationale and design of IMPACT

*BMC Medical Research Methodology, 2010;10:85*

### Chapter 3 Timely recruitment of the targeted number of participants perceived as more difficult than anticipated by many trialists - an analysis of a cohort of trials in the Netherlands Trial Register 28

### Chapter 4 Why do some centres recruit better than others? An analysis of recruitment rates in 17 randomized clinical trials in obstetrics and gynecology 44

### Chapter 5 Pregnant women’s concerns when invited to a randomized trial 62

### Chapter 6 Does recruitment for multicenter clinical trials improve dissemination and timely implementation of their results? A survey study from the Netherlands

*Clinical Trials 2013;10:915-23* 86
## PART 2: IMPROVING EVIDENCE-BASED CLINICAL PRACTICE

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Who teaches the Evidence-based Medicine teacher?</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td><em>Medical Teacher, 2012;34:866</em></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>How to confidently teach EBM on foot: development and evaluation of a web-based e-learning course</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td><em>Evidence-based Medicine, 2013;18:170-2</em></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>How can we teach EBM in clinical practice? An analysis of barriers to implementation of on-the-job EBM teaching and learning</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td><em>Medical Teacher, 2011;33:e125-30</em></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Tools to assess Evidence-based Practice behaviour among healthcare professionals.</td>
<td>134</td>
</tr>
<tr>
<td></td>
<td><em>Evidence-based Medicine, 2013;18:129-38</em></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Summary and general discussion</td>
<td>154</td>
</tr>
<tr>
<td></td>
<td>Samenvatting en algemene beschouwing</td>
<td>166</td>
</tr>
</tbody>
</table>

List of co-authors that contributed to this thesis | 180  |
PhD portfolio | 184  |
Dankwoord | 192  |
CHAPTER 1

GENERAL INTRODUCTION
Decisions need to be made continuously in every day clinical practice, either explicitly or more implicitly. Implicit or intuitive decisions are often based on knowledge gained by experience. It is well known that the way in which we derive knowledge from experience could lead to conclusions that are systematically biased.¹,²

One could imagine, for example, that an adverse event that happened to a patient after an intervention, stays more firmly etched in the physician’s memory than the hundred times the same intervention was performed before, without any complication occurring. After the physician has applied an alternative intervention dozens of times, without any adverse events occurring, he might intuitively conclude that the new, alternative intervention is safer. Is this wrong?

No, in itself, it’s not. It is perfectly human. If we were be able to appropriately mark everything that happens around us, we would be able to process a huge amount of information now mostly unused. What can be problematic, is an intervention judged as helpful, based on the collective experience of multiple experts, but then proven to be ineffective, or even harmful, when appropriately tested in systematic research.³,⁴

Overall, it seems unwise to rely exclusively on experience for important decisions in clinical medicine. New and existing interventions should be evaluated based on sound and solid trials of their effectiveness and potential harms. Clinical medicine should be a combination of personal wisdom and experience, amplified and corrected where needed by the rapidly growing body of evidence from scientific research. This integration of individual clinical expertise with the best available evidence and patients preferences, has been introduced as ‘Evidence-based Medicine’ (EBM) in 1992 by dr. Guyatt and his colleagues.⁵
This thesis consists of a number of research projects, centered on ‘evidence-based medicine’. We concentrated both on the generation of evidence, through clinical trials, and the integration of evidence from solid research into clinical practice.

Part 1 focuses on improving enrollment of patients in clinical trials, as this is a major problem encountered in the generation of scientific evidence. Part 2 focuses on the integration of evidence-based decision making in everyday clinical practice, by improving the teaching and evaluation of evidence-based clinical practice.

**PART 1: IMPROVING ENROLLMENT OF PATIENTS IN CLINICAL TRIALS**

The randomized trial is worldwide considered as the best instrument to evaluate the effectiveness of medical interventions, and, as such, it is a cornerstone in EBM. Optimal and affordable care, both from an individual and from a societal perspective, requires valid and reliable randomized trials, recruiting a sufficient number of participants, to generate the evidence for medical decision-making.

In many randomized trials recruitment is usually slower and more complex than expected. In a cohort of 114 multicenter trials funded by the UK Medical Research Council and the UK Health Technology Assessment Program between 1994 and 2002, less than one-third recruited their original target within the time originally planned, and around one third had extensions.\(^6\)

If the targeted sample size is not achieved, the clinical trial will have less statistical power to convincingly estimate the effectiveness of the medical intervention under study with sufficient precision, which could lead to erroneous healthcare decisions, or suboptimal dissemination of study findings. Reasons for slow recruitment can be found at different levels: a patient may decide not to participate, a clinician can refrain from inviting an eligible patient, and a department or hospital board could decide not to participate. Multiple reviews on barriers and motivators for participation in clinical trials have been published,\(^7\)-\(^{13}\) but it remains uncertain to what extent these results can be generalized to other trials and patient populations. Selection cannot be excluded, with an overrepresentation of either very successful or very unsuccessful recruiting trials in published evaluations of recruitment, as these trials can especially invoke a study of determinants of recruitment.
In *Chapter 2* we present the design of the IMPACT study - Improving Participation in Clinical Trials, in which we tried to capture characteristics associated with successful patient recruitment at three levels: the level of the trial organization (Chapter 3), the level of the doctor or hospital (Chapter 4) and the level of the patient (Chapter 5).

In the project reported in *Chapter 3* we studied recruitment at the trial level. A questionnaire regarding achieved recruitment and factors potentially influencing recruitment was sent to principal investigators of a cohort of trials registered in the Netherlands Trial Register. We looked at whether recruitment had been successful (at least 80% of the patients was recruited within the timeframe) and at characteristics associated with recruitment success.

In the study presented in *Chapter 4* we looked at recruitment at the level of the center. We sent a questionnaire to local coordinators of 17 trials running within the Dutch Consortium for Women’s health and Reproductivity studies, to look at factors influencing recruitment in their center. We looked at factors motivating centre decisions about participation in new trials, the research orientation of the department, and the (perceived) logistic support by research personnel.

*Chapter 5* presents an analysis of recruitment at the level of the patient. We performed semi-structured interviews with 21 women invited to participate in a trial in obstetrics. We invited in a 1:1 ratio patients who had participated and patients who had declined the invitation to be enrolled in one of eight trials in obstetrics: the Allo, Apostel I, Apostel II, Chips, WOMB, Ppromexil, Hypitat II and the ProTwin trial.

In *Chapter 6* we look at whether recruitment for clinical trials improves dissemination and timely implementation of the trial results. We sent a questionnaire to gynaecologists, residents, nurses and midwives in all centres in Obstetrics and Gynaecology in the Netherlands. For nine trials we asked whether they were aware of the trial results, were convinced by the results, and what percentage of their patients were treated according to the results of these trials. We compared the answers to these questions between respondents who worked in a hospital that had recruited for a trial and respondents who worked in a hospital that had not recruited for a trial.
PART 2: IMPROVING THE TEACHING AND EVALUATION OF EVIDENCE-BASED CLINICAL PRACTICE

A famous quote of Dr. Sackett is: ‘Half of what you’ll learn in medical school will be shown to be either dead wrong or out of date within five years of your graduation; the trouble is that nobody can tell you which half – so the most important thing to learn is how to learn on your own.’\textsuperscript{14,15} Given this fast expansion of medical knowledge, varying in quality and applicability on a patient, an explicit and systematic way to search for answers in daily clinical practice is needed: evidence-based medicine (EBM). Generally, the process of EBM is divided in 5 steps:\textsuperscript{5}

1. Define a clinically relevant question
2. Search for the best evidence
3. Critically appraise the evidence
4. Apply the evidence
5. Evaluate the performance of EBM

Although EBM is nowadays mostly considered an integral part of medical training, integration of (teaching) various steps in busy clinical practice remains a challenge. EBM teaching is an essential element for bringing EBM in clinical practice.

In the study reported in Chapter 7 we identified the availability and contents of Teach the Teacher EBM courses in various European countries. As a limited number of Teach the Teacher EBM courses was available, we describe the development and evaluation of a Teach the Teacher EBM e-learning course in Chapter 8. The e-learning course demonstrates in different clinical settings the application of EBM and aims to encourage integration of EBM teaching in daily clinical practice.

In Chapter 9 we look at barriers that refrain clinical teachers from teaching EBM in clinical practice in various European countries. We compared the level of the barriers between countries.

As there is currently no standard for evaluating how evidence-based individual health professionals or departments work, Chapter 10 presents a review of the literature for tools to assess the level of evidence-based practice performance of health professionals.
REFERENCES

7. Rendell JM, Merrit RK, Geddes J. Incentives and disincentives to participation in randomized controlled trials Cochrane Database of Systematic Reviews 2007. MR000021.
PART I

IMPROVING ENROLLMENT OF PATIENTS IN CLINICAL TRIALS
CHAPTER 2

IMPROVING PARTICIPATION OF PATIENTS IN CLINICAL TRIALS
- RATIONALE AND DESIGN OF IMPACT

Katrien Oude Rengerink, Brent C. Opmeer, Sabine L.M. Logtenberg, Lotty Hooft, Kitty W.M. Bloemenkamp, Monique C. Haak, Martijn A. Oudijk, Marc E. Spaanderman, Johannes J. Duvekot, Christine Willekes, Maria G. van Pampus, Martina M. Porath, Jim van Eyck, Marko J. Sikkema, Ben Willem J. Mol

BMC Medical Research Methodology, 2010;10:85
ABSTRACT

**Background** One of the most commonly reported problems of randomised trials is that recruitment is usually slower than expected. Trials will cost more and take longer, thus delaying the use of the results in clinical practice, and incomplete samples imply decreased statistical power and usefulness of its results. We aim to identify barriers and facilitators for successful patient recruitment at the level of the patient, the doctor and the hospital organization as well as the organization and design of trials over a broad range of studies.

**Methods/design** We will perform two cohort studies and a case-control study in the Netherlands. The first cohort study will report on a series of multicentre trials performed in a nationwide network of clinical trials in obstetrics and gynaecology. A questionnaire will be sent to all clinicians recruiting for these trials to identify determinants - aggregated at centre level - for the recruitment rate. In a case control-study nested in this cohort we will interview patients who refused or consented participation to identify factors associated with patients’ consent or refusal. In a second cohort study, we will study trials that were prospectively registered in the Netherlands Trial Register. Using a questionnaire survey we will assess whether issues on hospital organization, trial organization, planning and trial design were associated with successful recruitment, i.e. 80% of the predefined number of patients recruited within the planned time.

**Discussion** This study will provide insight in barriers and facilitators for successful patient recruitment in trials. The results will be used to provide recommendations and a checklist for individual trialists to identify potential pitfalls for recruitment and judge the feasibility prior to the start of the study. Identified barriers and motivators coupled to evidence-based interventions can improve recruitment of patients in clinical trials.
BACKGROUND

Evaluation research is essential to inform evidence based health care decisions. The randomized controlled trial (RCT) is worldwide considered as the best instrument to evaluate the effectiveness of medical interventions. One of the most commonly reported problems with the conduct of such RCTs, however, is that recruitment is usually slower than expected. In the 1970’s an American pharmacologist, Luis Lasagna, stated that once trial recruitment starts, the supply of eligible patients becomes a fraction of what was assumed before the start of the trial. This phenomenon, currently known as Lasagna’s Law, still holds today.\textsuperscript{1,2} In the UK, in a cohort of 114 multicentre trials funded by the UK Medical Research Council and the UK Health Technology Assessment Programme between 1994 and 2002, less than one-third recruited their original target within the time originally planned, and around one third had extensions.\textsuperscript{1}

If in a clinical trial the targeted sample size is not achieved, it will have less statistical power to convincingly demonstrate potentially important differences between the groups, which might make the results less useful or not at all applicable in clinical practice – it will not improve practice and wastes the contribution of participants who already participated. In addition, if recruitment has to be extended to reach the required sample size, the trial will cost more and take longer, thus delaying the use of its results in clinical practice. As the total amount of funding is limited, fewer trials can be conducted and hence less clinical dilemmas can be solved.

Figure 1: Factors for lack of recruitment can be found on different levels: with interactions between levels.

<table>
<thead>
<tr>
<th>Study Level</th>
<th>Did the study recruit well?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Center Level</td>
<td>If the study is recruiting, did the center recruit well?</td>
</tr>
<tr>
<td>Doctor Level</td>
<td>If center is participating, did doctor recruit well?</td>
</tr>
<tr>
<td>Patient Level</td>
<td>If the patient is asked, did he/she consent?</td>
</tr>
</tbody>
</table>
Reasons for lack of recruitment can be found at different levels: the patient, the doctor, the participating centre or department, the study organisation and the study design (see figure 1). Rendell et al. reviewed incentives and disincentives to participation, focusing on participation of clinicians. Lovato et al. and Ross et al. reviewed barriers to both patient and clinician participation. Other reviews were dedicated to a specific disease or group of diseases.

Frequently mentioned barriers for patients are preference for one form of treatment, concerns with the trial setting, dislike of randomisation, general discomfort with the research process, distrust in researchers, complexity and stringency of the protocol, presence of a placebo or no-treatment group, potential side effect, fear that trial involvement would have a negative effect on the relationship with their physician and their physicians’ attitudes towards the trial, the potential for increased demands and the mere inability to make a decision. Frequently mentioned barriers to clinician participation are time constraints, lack of staff and training, worry about the impact on doctor-patient relationship, concern for patients, loss of professional autonomy, difficulty with the consent procedure, lack of rewards and recognition, and an insufficiently interesting question.

These reasons might be different across specialties, countries or due to the nature of the disease or disease population, but will probably also have common denominators. In many of the included papers in the reviews it is uncertain if and how these results can be generalized to other trials and populations. In the reviews there might also be an over representation of very successful or very unsuccessful recruiting trials, as especially these trials might invoke a study of determinants of recruitment by the trial coordinator.

Recently, the STEPS study focused on recruitment from different perspectives, and although they identified determinants, they did not have sufficiently definitive results to make strong recommendations.

In conclusion, although several potential barriers have been identified, it remains unclear whether they are applicable for any next trial. At present, neither a rule which can predict successful recruitment, nor a checklist to assess the feasibility of recruitment prior to the start of a trial, or even prior to a funding decision, is available. We therefore aim to study methods for IMproving PArticipation in Clinical Trials (IMPACT). We aim to identify predictors for (un)successful patient recruitment at the level of the patient, the doctor and the study and organizational level in order to provide a rule or checklist for successful patient recruitment in clinical trials.
METHODS/DESIGN

To address this topic at different relevant levels, we plan two cohort studies and a nested case control study in The Netherlands. In the first cohort study we will focus on successful recruitment at the level of participating centres, where characteristics of the clinician will be aggregated at the level of the centre. In a case control study embedded in this cohort study we will focus on the patient level. In the second cohort study we will identify determinants for success of recruitment in a cohort of trials registered in the national Netherlands Trial Register.

This study did not require formal approval of an ethics committee or internal review board, as was confirmed by the ethics committee of the Academic Medical Center. We interpreted completion and return of a questionnaire as the respondent’s consent for participation. Methodological details of these studies are described below.

COHORT STUDY 1: PREDICTING RECRUITMENT AT CENTER LEVEL

The first cohort aims to determine which factors on centre level will influence recruitment of patients. The cohort will consist of a series of multicenter trials performed in a nationwide Consortium on studies in Obstetrics and Gynaecology, in which currently over 70 medical centres participate (table 1). At the centre level, we will study aspects of the doctor as well as aspects of the organisation. A questionnaire will be sent to gynaecologists, residents, research nurses, and midwives in the Netherlands who work in centres recruiting for these trials. We will also collect data on characteristics of clinicians and study organization.

Based on these data, we will construct a prediction model. The primary outcome will be the percentage of eligible patients recruited, defined as the number of randomized patients per centre divided by the number of eligible patients. Secondary outcomes will be the number of randomized patients divided by the number of available patients per centre stratified per study. The number of eligible patients is registered for a part of the studies and will be estimated based on the LVR, the Dutch national perinatal registry.

As potential predictors we will take into account characteristics of the clinicians, e.g. the proportion of doctors with a PhD, as well as clinicians’ views on the trial, e.g. prior belief in the relevance and quality of the study design. We will also collect characteristics of the study organisation: e.g. status of the hospital (academic, teaching or general), availability
of research nurses or employees for counselling of the studies, clarity of research protocol and logistics, responsibilities for recruitment. These potential predictors reflect clinicians’ views on these topics, aggregated at the level of department.

For reliable modelling on prediction experts recommend that there should be at least about 10 events in the data set for each potential predictor to be included in the model.\(^{16}\)

As we expect a response in about 65 of 70 centres, prediction models can include about 6-7 predictors at centre level. Missing values will be imputed. The predictive accuracy of the model will be assessed using calibration, which evaluates the correspondence between the model’s predicted percentage of randomized patients and the observed percentage of randomized patients. The discriminative ability of the prognostic model will be assessed by using receiver operating characteristics (ROC) analysis.

We will then perform internal validation using bootstrapping and apply shrinkage to correct for over fit. A simplified prediction rule will be derived from the regression coefficients of the independent predictors in the multivariable model. If a valid prediction rule cannot be constructed, we will use the multivariable model to identify risk factors for a low recruitment percentage.

CASE CONTROL STUDY: PREDICTING RECRUITMENT AT PATIENT LEVEL

In a case control-study nested in this cohort we will interview patients who refused or consented participation in a set of clinical trials, to identify factors which influenced the decision to participate.

We will perform qualitative semi-structured interviews. The interviews will start with open questions on the motivation of a patient to consent or refusal participation. Subsequently, the interview will be guided by a topic list that is based on the literature and on input from experienced trialists. Topics will include counselling, clarity and understanding of patient information, knowledge of and attitude towards scientific research,\(^{9,17}\) attitude towards the doctor or health care organisation, type of intervention, practical considerations and organisational issues, (perceived) personal benefit, (dis)trust and their social demographic characteristics. This topic list will be tailored during the study.

We will interview patients who were recently counselled for participation in a RCT in a broad range of trials in the field of obstetrics, subfertility, gynaecology, internal medicine, neurology and surgery. Patients will be selected using purposive sampling. We will
Table 1: A cohort of studies performed in a nationwide Consortium on studies in obstetrics and gynaecology in the Netherlands.

<table>
<thead>
<tr>
<th>Nr</th>
<th>Trial</th>
<th>Population</th>
<th>Sample size</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Amphia</td>
<td>Women with a multiple pregnancy before 20 weeks pregnancy</td>
<td>720</td>
</tr>
<tr>
<td>2</td>
<td>Hypitat II</td>
<td>Women with pregnancy induced hypertension or mild pre-eclampsia at 34-37 weeks gestation</td>
<td>400</td>
</tr>
<tr>
<td>3</td>
<td>Digitat</td>
<td>Women with a singleton pregnancy at 36 completed weeks of gestation or more</td>
<td>626</td>
</tr>
<tr>
<td>4</td>
<td>Ppromexil</td>
<td>Women with preterm prelabour rupture of membranes between 34 and 37 weeks gestation</td>
<td>520</td>
</tr>
<tr>
<td>5</td>
<td>Stan</td>
<td>Women in labour over 36 weeks of gestation with an indication for CTG monitoring</td>
<td>2400</td>
</tr>
<tr>
<td>6</td>
<td>Apostel 1</td>
<td>Women with threatened preterm labour at 24-34 weeks gestation.</td>
<td>220</td>
</tr>
<tr>
<td>7</td>
<td>Apostel 2</td>
<td>Women with threatened preterm labour at 26-32+2 weeks gestation.</td>
<td>400</td>
</tr>
<tr>
<td>8</td>
<td>TRIPLE P</td>
<td>Women with a singleton pregnancy</td>
<td>1920</td>
</tr>
<tr>
<td>9</td>
<td>Probaat</td>
<td>Women ≥ 37 weeks of gestation and Bishop score &lt; 6</td>
<td>812</td>
</tr>
<tr>
<td>10</td>
<td>PreCare</td>
<td>Women with preeclampsia or HELLP in previous pregnancy</td>
<td>250</td>
</tr>
<tr>
<td>11</td>
<td>WOMB</td>
<td>Women with &gt;1000 mL postpartum fluxus</td>
<td>400</td>
</tr>
<tr>
<td>12</td>
<td>Truffle</td>
<td>Women at 26-32 weeks gestation with a fetus with intrauterine growth retardation.</td>
<td>500</td>
</tr>
<tr>
<td>13</td>
<td>Allo</td>
<td>Women with suspected fetal asphyxia during labour</td>
<td>220</td>
</tr>
<tr>
<td>14</td>
<td>ProTWIN</td>
<td>Women with a multiple pregnancy between 12 and 20 weeks gestation</td>
<td>660</td>
</tr>
<tr>
<td>15</td>
<td>Metex</td>
<td>Women with an extra uterine gravidity</td>
<td>72</td>
</tr>
<tr>
<td>16</td>
<td>ESEP</td>
<td>Women with an extra uterine gravidity in one of the tubae and a normal contralateral tuba</td>
<td>450</td>
</tr>
<tr>
<td>17</td>
<td>INeS</td>
<td>Couples with unexplained subfertility or a mild male factor</td>
<td>600</td>
</tr>
<tr>
<td>19</td>
<td>MOVIN’</td>
<td>Women with anovulation not pregnant after 6 ovulatory cycles of clomid</td>
<td>200</td>
</tr>
<tr>
<td>20</td>
<td>Bedrest</td>
<td>Women who undergo intra-uterine insemination</td>
<td>250</td>
</tr>
<tr>
<td>21</td>
<td>VUSIS 1</td>
<td>Women with stress-incontinence</td>
<td>100</td>
</tr>
<tr>
<td>22</td>
<td>PORTRET</td>
<td>Women with stress incontinence aged 35-80 years</td>
<td>400</td>
</tr>
<tr>
<td>23</td>
<td>CUPIDO</td>
<td>Women with vaginal prolaps</td>
<td>114</td>
</tr>
<tr>
<td>24</td>
<td>Pompoen</td>
<td>Women with postmenopausal bleeding</td>
<td>200</td>
</tr>
</tbody>
</table>

Interview patients from different doctors, different natures of disease, different levels of education, and different regions in the Netherlands.
The number of patients to be interviewed will be dependent on the variety of responses until saturation of the data is reached. We expect this to be about 10 patients who participated and 10 patients who refused for the first subspecialty obstetrics. As we assume that reasons for the decision about participation share common components across specialties, we can complete the reasons with a smaller sample of patients from other clinical specialties. This will provide us information about the full spectrum of barriers and facilitators for participation.

Based on these barriers and motivators observed in the interviews we will construct a questionnaire to quantify these findings in a representative sample of patients in the fields of obstetrics and gynaecology, neurology, surgery and internal medicine.

COHORT STUDY 2: PREDICTING RECRUITMENT SUCCESS OF TRIALS REGISTERED IN THE NETHERLANDS TRIAL REGISTER

In a second cohort study, we will include trials that have been registered prospectively in the Netherlands Trial Register. The cohort of studies will consist of all studies that registered their stop date between January 1st 2005 and January 1st, 2010 (expected number of available trials N≈1000). Using a questionnaire survey, we will investigate whether issues on hospital organization, trial organization, planning and trial design are predictive for successful recruitment, defined as ≥80% of the patients recruited within the time frame defined in the grant application.

As potential predictors we will take into account trial characteristics, i.e. placebo arm, blinding, experimental status of the intervention to be evaluated; characteristics of the trial organisation, i.e. research staff available to counsel patients and acquire follow up data, who is responsible for recruitment; and characteristics of the principal investigator and the research group, i.e. composition of different expertise, experience and training in trial research. For reliable prediction modelling experts recommend that there should be at least about 10 events in the data set for each potential predictor to be included in the model. In a sample of 1000 studies, with a (low) recruitment rate of 30% about 30 potential predictors can be tested reliably.

Missing values will be imputed. Like in the cohort study I, the predictive accuracy of the model will be assessed using calibration, which evaluates the correspondence between the model’s predicted probabilities of recruitment success and the observed recruitment success over groups. The discriminative ability of the prognostic model will be assessed by
using receiver operating characteristics (ROC) analysis. We will perform internal validation using bootstrapping and apply shrinkage to correct for over fit.

A prediction rule will be derived from the regression coefficients of the independent predictors in the multivariable prognostic model. If a valid prediction rule cannot be constructed, we will use the model to identify risk factors for a low recruitment percentage.

DISCUSSION

At the end of the study, we will have an inventory of predictors, barriers and facilitators for successful patient recruitment in trials. The first cohort study will provide information on prediction of recruitment in different centres participating in obstetrical, subfertility and gynaecological trials. From the nested case-control study we will obtain qualitative and quantitative information about factors which influence patient participation. The second cohort study will provide insight in factors related to successful recruitment at study level. Based on the prediction models or risk factors for unsuccessful recruitment, we will develop a set of recommendations and a checklist that can be used by individual trialists before the start of the study to assess if recruitment of the proposed sample size with their strategy will be feasible.

A strong point of the design is that we will address recruitment factors at different levels and from different perspectives, throughout a variety of trials in various fields of medicine. This will provide us with a broad overall picture of reasons why patients participate or refuse participation, and will provide insight in a common denominator between trials, or clarify differences. At the same time this is a potential pitfall: the reasons for participation or non-participation in a clinical trial might predominantly depend on exclusive characteristics of a trial and its targeted population, so that general predictors may not be identified. If so, the results are still valuable, but we should focus on the development of a general applicable recruitment tool. Such a tool might consist of a strategy based on interviewing a number of eligible participants as well as a number of recruiters and/or clinicians prior to start or during the piloting of the trial. Another strong point is that the first cohort is based on clinicians recruiting for a set of
trials from the nationwide consortium in Obstetrics and Gynaecology. All academic medical centres and the majority of the Dutch hospitals recruit for trials running in this consortium, which enables us to cover a large part of the Netherlands without selection of the explicitly research minded hospitals. Moreover, the second cohort of studies from the Dutch Trial Register will be a representative sample of all trials performed in the Netherlands, since from July 2005 registration is required for publication in important journals. It will be a challenge to deliberately handle the heterogeneity of the trials included.

There are also some limitations in this study which require a remark. First, although we focus on different levels (level of the patient, centre and study) it is not possible to directly link the data from these levels. We therefore cannot disentangle the relative impact of each level to the recruitment problems. However, we will be able to provide a satisfactory estimation, given the variability between different specialties and trials.

Second, as we focus on trials in obstetrics and gynaecology in the first cohort study, generalizability of these results to other clinical specialties should be evaluated. As in this study it is not feasible to survey all specialties in depth, we think it is more informative to have a complete picture of one specialty over a limited amount of information from many specialties.

Furthermore, in the second cohort on predicting successful recruitment on study level, the predefined recruitment target will probably - as most power calculations - be based on limited information. When information emerges from external sources or interim analyses, the sample size might be adapted, which can make the trial more successful or efficient without reaching the originally planned sample size. However, sample size targets are taken into account when funding decisions are made. The extent to which a trial meets initial expectations can be viewed as a legitimate marker of trial success. Moreover, recruitment can be viewed of as a surrogate marker of more significant markers of success, such as the extent to which the trial question has been successfully addressed.¹

Interviewing patients from different specialties and trials will provide a broad spectrum of why patients participate or refuse. This will limit the number of patients interviewed from one trial, but the quantification of these reasons in a representative group of patients will allow us to examine differences between these specialties and trials.

Finally, as this study will be performed in the Netherlands, the Dutch health care system as well as the position of medical research in the Netherlands might influence its results.
Existing literature can be used to compare our results with those from other countries. We realize that especially to make this study a success we need a high response rate to avoid selection of highly motivated researchers, patients and clinicians.

In conclusion, this design allows us to identify determinants for unsuccessful recruitment. Identified predictors for unsuccessful recruitment can be coupled to evidence-based strategies to improve recruitment in trials.

Acknowledgements
This study is funded by ZonMw Grant 80-82310-97-10210. We thank three anonymous reviewers for commenting on the initial study design.
REFERENCES

CHAPTER 3
TIMELY RECRUITMENT OF THE TARGETED NUMBER OF PARTICIPANTS PERCEIVED AS MORE DIFFICULT THAN ANTICIPATED BY MANY TRIALISTS
AN ANALYSIS OF A COHORT OF TRIALS IN THE NETHERLANDS TRIAL REGISTER
Katrien Oude Rengerink, Patrick M.M. Bossuyt, Ben Willem J. Mol, Lotty Hooft
Chapter 3 | Timely recruitment of the targeted number of participants

ABSTRACT

Introduction Patient recruitment in clinical trials often takes longer than expected. Trials with slow recruitment are more costly and an insufficient sample size can lead to indecisive conclusions.

Design We sent a questionnaire to principal investigators of 1,129 prospectively registered trials in the Netherlands Trial Register with an expected date of completing recruitment between 2005 and 2010. We analysed recruitment success and recruitment duration in these trials, classifying them by the proportion recruited at the planned completion date relative to the target number, as unsuccessful (<80% recruited) or successful (≥80%). We additionally evaluated whether or not the target number had (≥100%) or had not been reached at the actual stop date (<100%). In case of unsuccessful recruitment we asked the principal investigator to indicate why recruitment had been suboptimal, and how recruitment could be improved in the future. We assessed whether characteristics of the trial design, organisation or coordinator were associated with successful recruitment, with a focus on factors than can be modified.

Results We received 386 complete questionnaires (34%). In 165 trials (43%) 80% or more of the patients had been enrolled at the planned completion date; 183 trials (47%) had recruited fewer than the targeted number at the actual stop date. In 11% of trials inclusion criteria were modified to facilitate patient recruitment. The median actual duration of patient recruitment was 18 months, while median 12 months recruitment was planned. Of the trial coordinators 244 (66%) stated that recruitment had been more difficult than anticipated. They mentioned the following reasons most frequently: missing eligible patients (29%), fewer patients eligible than expected (29%), more patients than expected declined the invitation to participate (24%). A consistent pattern of characteristics of the trial design, organisation or coordinator associated with recruitment success could not be identified. Trialists mentioned the following potential solutions to facilitate recruitment: better information to potential participants, appointing a (dedicated) research nurse, positive publicity about trials to both the general public as clinicians, performing multicentre studies, and more collaboration between researchers, and recruiters with more dedication to research.

Conclusion The majority of trialists is initially too optimistic about the number of patients participating in trials. It is unlikely that there is a simple series of trial features consistently associated with recruitment success; a multifaceted approach is probably needed to facilitate enrolment of the targeted number of participants in future trials.
INTRODUCTION

The randomized controlled trial (RCT) is worldwide considered the best instrument to evaluate the effectiveness of medical interventions. One of the most commonly reported problems in RTC’s is slower than expected recruitment of participants, and in many RCT the targeted sample size is never reached. An American pharmacologist, Luis Lasagna, stated in the 1970s that, when trial recruitment starts, the supply of suitable patients becomes a fraction of what was assumed before the start of the trial. This phenomenon, currently known as Lasagna’s Law, still seems to hold today.\(^1,2\) In a cohort of 114 multicentre trials funded by the UK Medical Research Council and the UK Health Technology Assessment Programme between 1994 and 2002, less than one-third had recruited their original target within the time originally planned, and around one third needed extensions of the recruitment period.\(^1\)

Failure to reach the targeted sample size in a trial means that the study will lack the precision in estimating effectiveness, which might hamper decision making based on the trial results. Overestimating the speed of enrolment could invite modifications of the trial protocol, making the study potentially more vulnerable, and can invoke extensions, which makes trials typically more costly and postpones the analysis.

Reasons for slow recruitment have been identified at different levels: at the level of patients, health professionals, and the study sites. A number of reviews on barriers and facilitators have been performed, focusing on both patient and clinician participation in general,\(^3-5\) or dedicated to a specific disease, or group of diseases.\(^6-11\) Treweek et al. reviewed strategies to improve recruitment in clinical trials, tested in a randomized or quasi-randomized trial, including those recruiting to hypothetical studies. They found that telephone reminders, opt-out instead of opt-in for being contacted by the trial team regarding participation, and open designs improved trial participation.\(^12\) The effect of many other interventions to improve patient enrolment was less clear. Individual studies have looked at trial characteristics that might influence recruitment, such as (understanding of) randomisation, having a placebo arm, blinding, the prevalence of the disease, and eligibility criteria.\(^13-16\) Most of these previous analyses were based on selected trials, which hampers generalization. An exception is the STEPS study, in which Campbell and colleagues identified determinants in an unselected series of trials, but their sample size was
We set out to identify characteristics of the trial design, organisation or coordinator associated with recruitment success in a large cohort of consecutive clinical trials registered in the Netherlands Trial Register.

METHODS

DESIGN
We evaluated recruitment in a cohort of trials registered in the Netherlands Trial Register (NTR), looking at potential determinants (characteristics of the trial design, organisation or coordinator) of successful recruitment using a questionnaire survey. This analysis is part of the IMPACT project, of which the full protocol is available elsewhere. According to Dutch law, our study did not require informed consent or IRB approval, confirmed by the executive board of the medical ethics committee of the Academic Medical Centre of the University of Amsterdam (#10.17.1463).

COHORT OF TRIALS
We identified all trials that had been registered before study initiation in the NTR with a planned completion date between January 1st 2005 and January 1st 2010. Eligible were trials in which coordination was based in the Netherlands. The planned completion date was selected as a key inclusion criterion, which made that data on achieved recruitment was collected at least 18 months after the planned completion. Choosing the stop date instead of the start date also allowed inclusion of trials with a longer planned recruitment period. No additional exclusion criteria were used.

DATA COLLECTION AND QUESTIONNAIRE
From the NTR we extracted key trial characteristics. For additional information, in June 2011 we sent a link to an online questionnaire by email to the trial coordinators, using contact information in the registry. In the accompanying message we asked that the person best informed about trial coordination, logistics and recruitment would fill out the questionnaire. If the email address was no longer in use, we searched Google or PubMed for a more recent address. If we did not receive a response to the initial email, two or more reminders were sent, and a ‘happy new year’ card as a questionnaire reminder. If all of this did not result in a response we tried to contact the researchers by phone, using multiple attempts.
The questionnaire was set in Dutch and consisted of five sections: 1) recruitment characteristics, 2) trial characteristics, 3) trial organisation, 4) trial registration, and 5) principal investigator or trial coordinator. We also asked how trialists thought recruitment could be improved. The potential determinants for successful recruitment were selected based on a systematic search of reviews about factors potentially influencing recruitment, articles about factors influencing recruitment, published questionnaires constructed for the identification of factors influencing recruitment in trials, and personal experience in the study group. We interpreted completion and return of the questionnaire as the respondent’s consent for participation. The full questionnaire can be received from the author.

OUTCOMES
We analysed recruitment success and recruitment duration in the included trials, classifying them by the proportion recruited at the planned completion date relative to the target number, as unsuccessful (<80% recruited) or successful (≥80%). We additionally evaluated whether or not the target number had (≥100%) or had not been reached at the actual stop date (<100%), after an eventual extension period.

STUDY REPORT
We evaluated failure to publish trial results, by searching Medline (through PubMed) in January 2012 for a study report, published two or more (up to seven) years after the planned completion date. To identify the correct article, we used the trial registration number, single entries or combinations of the registry record title, name of the contact author, intervention and/or primary outcome using Boolean operators AND/OR.

ANALYSIS
We summarized responses to the survey based on all completed and returned questionnaires. The percentage of trials with successful recruitment and the percentage with completed recruitment were calculated. We used chi-square test statistics to evaluate potential determinants of recruitment success in univariable analysis (listed in Table 1). In these analyses, we excluded trials that had never started recruitment for whatever reason: for example, because they never got ethical approval, or because guidelines had changed before start of the trial, making the trial redundant. In addition, we calculated the number of completed trials for which a study report was available in the literature. In our analyses a significance level of 0.05 was used.
RESULTS

RESPONSE AND DESCRIPTION OF TRIALS
The questionnaire was returned for 386 trials (34%). In total, these trials had aimed to recruit 242,412 participants, with a median targeted sample size of 120 (IQR 50 to 300; minimum 6, maximum 90,000).

RECRUITMENT SUCCESS
Recruitment had been successful (≥80% recruited) in 165 of the 386 trials (43%). At the actual stop date 183 (47%) had incomplete recruitment recruitment (fewer patients than targeted), even after an extension in some cases. Two thirds of the trialists (244 of 368; 66%) stated that recruitment had been more difficult than anticipated. The median duration of planned recruitment was 12 months; the median actual enrolment period was 18 months.

FACTORS ASSOCIATED WITH RECRUITMENT
We excluded 22 trials from our analyses, because they had never started recruitment, leaving 364 trials in our dataset. Table 1a shows associations between trial characteristics and successful and completed recruitment, table 1b shows associations between trial characteristics and successful recruitment restricted to multicentre trials, and table 1c focuses on factors that could be modified in a next trial. Below we present our analyses for a number of subgroups.

1. Randomised versus non-randomised trials
In 86% of the trials for which a questionnaire was returned randomized was performed (Table 1a). Recruitment had been successful in 44% of these, versus in 54% of non-randomised trials (RR 0.81, 95% CI 0.58-1.13). Successful recruitment was more often observed in trials where randomisation was done with envelopes (49%) or by telephone (46%), compared to web-based randomisation (34%), but this difference was not statistically significant.

2. Multicentre and single centre trials
On average, single centre studies recruited significantly more often successfully than multicentre studies: 80 of 161 (50%) versus 69 of 188 (37%) (RR 1.35; 95% CI 1.06-1.73, p=0.014). This difference was smaller when the actual stop date of recruitment was taken into account: 55% for single centre studies versus 45% in multicentre trials. The
targeted sample size was larger in multicentre trials: a median of 183 (IQR 93-434) versus 75 in single centre studies (IQR 25-183). As shown in Table 1b, paid research staff was not significantly associated with recruitment success, neither were participation of all centres in the grant application, a financial incentive for inclusion, or agreement about co-authorship before the start of the trial.

3. Recruitment success for the top specialties in number of trials performed

Most trials had been started in the field of internal medicine, where recruitment was successful in 46% of 48 trials (Table 1a). Lower proportions of successful recruitment were observed in psychiatry (6 of 25; 25%) and in general practice (7 of 22; 32%); when looking at completed recruitment these were no longer apparent.

4. Modifiable factors

We studied whether factors that could be modified in a subsequent trial were associated
with recruitment (Table 1c). There was no significant difference in recruitment success between trials with or without a feasibility or pilot study before the start of the trial: 62 of 126 (49%) for those with versus 98 of 227 for those without (43%). Involvement of patients in the definition of primary outcomes before starting the trial and/or exploring patients’ interest in the study were negatively associated with recruitment success (41% versus 46%).

Recruitment was also less frequently successful in trials in which a presentation was given to recruiters before start of the trial, in trials in which trial pockets cards were used, and in those in which an email was sent at the start of a study.

Recruitment was more often successful in trials in which responsibilities for recruitment were seen as very clearly defined. On a 5 point Likert scale we asked the trial coordinators how responsible they felt for the recruited number of patients in the trial, with 1 “not at all responsible” to 5 “very responsible”. If the trial coordinator stated that he or she felt

Table 1b: Trial characteristics and recruitment success, applicable to multicenter trials

<table>
<thead>
<tr>
<th>Trial characteristics</th>
<th>Patients recruited within planned timeframe</th>
<th>Patients recruited at actual stop date</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;80% (%)</td>
<td>≥80% (%)</td>
</tr>
<tr>
<td>Participation of all centres in grant application</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>73 (61%)</td>
<td>46 (39%)</td>
</tr>
<tr>
<td>No</td>
<td>49 (64%)</td>
<td>23 (37%)</td>
</tr>
<tr>
<td>Financial incentive for inclusion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>93 (63%)</td>
<td>55 (37%)</td>
</tr>
<tr>
<td>No</td>
<td>21 (60%)</td>
<td>14 (40%)</td>
</tr>
<tr>
<td>Recruiters co-authorship</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes, depends on recruitment</td>
<td>21 (54%)</td>
<td>18 (46%)</td>
</tr>
<tr>
<td>No</td>
<td>12 (71%)</td>
<td>5 (29%)</td>
</tr>
<tr>
<td>Unclear</td>
<td>70 (66%)</td>
<td>36 (34%)</td>
</tr>
<tr>
<td>Recruiters co-authorship</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Recruitment multiple countries</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Netherlands only</td>
<td>99 (62%)</td>
<td>62 (39%)</td>
</tr>
<tr>
<td>Yes at least one other country</td>
<td>11 (69%)</td>
<td>5 (31%)</td>
</tr>
<tr>
<td>Yes &gt;5 countries</td>
<td>6 (100%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Paid research staff available</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes in all centers</td>
<td>29 (73%)</td>
<td>11 (28%)</td>
</tr>
<tr>
<td>Only in large centers</td>
<td>4 (40%)</td>
<td>6 (60%)</td>
</tr>
<tr>
<td>Only in some centers</td>
<td>18 (62%)</td>
<td>11 (38%)</td>
</tr>
<tr>
<td>No</td>
<td>63 (61%)</td>
<td>41 (39%)</td>
</tr>
<tr>
<td>If so, what was their task?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Identification of patients</td>
<td>28 (67%)</td>
<td>14 (33%)</td>
</tr>
<tr>
<td>Other</td>
<td>91 (62%)</td>
<td>55 (38%)</td>
</tr>
</tbody>
</table>
“very responsible”, for the number of patients recruited in the trial, a higher percentage of trials reached their recruitment target (116 of 194 trials, 60%), compared with trials run by coordinators who stated a responsibility less than 5 “very responsible” (55 of 161 trials, 47%).

POTENTIAL SOLUTIONS FOR RECRUITMENT DIFFICULTIES
In response to a multiple choice question (multiple answers allowed) on the most important reason for recruitment difficulties the following factors were selected: eligible patients were missed (29%), fewer people were eligible than expected (29%), many eligible patients declined the invitation to participate (24%), internal problems (12%), such as lack of research staff and external problems (5%), like lack of publicity. In 25% of cases ‘another reason’ was selected, such as delays related to ethical review, or fewer centres than anticipated recruiting for a multicentre trial.

As a potential remedy, inclusion criteria were adjusted to facilitate patient recruitment in 38 trials (11%; in 4 trials (1%) for other reasons). The targeted sample size was adjusted in 80 trials (22%), related or unrelated to recruitment difficulties.

In response to the open question what trial coordinators perceived as the best strategy to facilitate recruitment a wide variety of strategies was mentioned (multiple options could be provided by one respondent). Trialists often mentioned providing clear (or better) information to patients, in a no-rush atmosphere (n=33), for which training of recruiters could be helpful (n=8). Dedicated research staff or research nurses (n=22), located in the different recruitment sites, was seen as helpful, both for identifying patients and for counselling. A higher reimbursement or incentive for recruiters (n=20) and participants (n=17), more multicentre, instead of single centre trials (n=17), more publicity about the study (n=19), a trial coordinator who frequently contacts recruiters and is available for any questions were also viewed as positively influencing recruitment. More collaboration was also seen as necessary to increase recruitment (n=12). Several respondents mentioned that nowadays trials do not always generate positive news in the media, a trend that they felt should be curbed (n=15). They indicated that one should educate the general public about the necessity of trials, presenting trials as an essential part of health care, instead of only experimental, removing unfortunate associations with mice and guinea pigs. Trialists also mentioned that estimations of their expected recruitment are (whether or not deliberately) too optimistic (n=22).
Table 1c: Factors that could be modified in a next trial

<table>
<thead>
<tr>
<th>Characteristic trial organisation and coordinator</th>
<th>Patients recruited within planned timeframe</th>
<th>Patients recruited at the actual stop date</th>
<th>p-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&lt;80%</td>
<td>≥80%</td>
<td>&lt;100%</td>
<td>≥100%</td>
</tr>
<tr>
<td>Pilot study</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>64 (51%)</td>
<td>62 (49%)</td>
<td>57 (45%)</td>
<td>69 (55%)</td>
</tr>
<tr>
<td>No</td>
<td>129 (57%)</td>
<td>98 (43%)</td>
<td>101 (45%)</td>
<td>126 (56%)</td>
</tr>
<tr>
<td>Patient involvement</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>27 (59%)</td>
<td>19 (41%)</td>
<td>26 (57%)</td>
<td>20 (44%)</td>
</tr>
<tr>
<td>No</td>
<td>176 (54%)</td>
<td>151 (46%)</td>
<td>143 (44%)</td>
<td>184 (56%)</td>
</tr>
<tr>
<td>Presentation recruiters</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>137 (62%)</td>
<td>85 (38%)</td>
<td>109 (49%)</td>
<td>113 (51%)</td>
</tr>
<tr>
<td>No</td>
<td>68 (48%)</td>
<td>74 (52%)</td>
<td>65 (46%)</td>
<td>77 (54%)</td>
</tr>
<tr>
<td>Presentation patients</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>27 (66%)</td>
<td>14 (34%)</td>
<td>20 (49%)</td>
<td>21 (51%)</td>
</tr>
<tr>
<td>No</td>
<td>178 (55%)</td>
<td>145 (45%)</td>
<td>154 (48%)</td>
<td>169 (52%)</td>
</tr>
<tr>
<td>Newsletter participants</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>22 (55%)</td>
<td>18 (45%)</td>
<td>16 (40%)</td>
<td>24 (60%)</td>
</tr>
<tr>
<td>No</td>
<td>183 (57%)</td>
<td>141 (44%)</td>
<td>158 (49%)</td>
<td>166 (51%)</td>
</tr>
<tr>
<td>Trial pocket cards</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>60 (65%)</td>
<td>32 (35%)</td>
<td>49 (53%)</td>
<td>43 (47%)</td>
</tr>
<tr>
<td>No</td>
<td>145 (53%)</td>
<td>127 (47%)</td>
<td>125 (49%)</td>
<td>147 (54%)</td>
</tr>
<tr>
<td>Start announced at website</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>56 (57%)</td>
<td>42 (43%)</td>
<td>41 (42%)</td>
<td>57 (58%)</td>
</tr>
<tr>
<td>No</td>
<td>149 (56%)</td>
<td>117 (44%)</td>
<td>133 (50%)</td>
<td>133 (50%)</td>
</tr>
<tr>
<td>Email at start</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>81 (68%)</td>
<td>38 (32%)</td>
<td>63 (53%)</td>
<td>56 (47%)</td>
</tr>
<tr>
<td>No</td>
<td>124 (51%)</td>
<td>121 (49%)</td>
<td>111 (45%)</td>
<td>134 (55%)</td>
</tr>
<tr>
<td>Announcement media</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>40 (50%)</td>
<td>40 (50%)</td>
<td>39 (49%)</td>
<td>41 (51%)</td>
</tr>
<tr>
<td>No</td>
<td>165 (58%)</td>
<td>119 (42%)</td>
<td>135 (48%)</td>
<td>149 (53%)</td>
</tr>
<tr>
<td>Extra visit necessary</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>45 (61%)</td>
<td>29 (39%)</td>
<td>33 (45%)</td>
<td>41 (55%)</td>
</tr>
<tr>
<td>No</td>
<td>160 (55%)</td>
<td>130 (45%)</td>
<td>141 (49%)</td>
<td>149 (51%)</td>
</tr>
<tr>
<td>Responsibility recruitment clear</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Very clear</td>
<td>141 (53%)</td>
<td>124 (47%)</td>
<td>117 (44%)</td>
<td>148 (56%)</td>
</tr>
<tr>
<td>Mostly clear or not very clear</td>
<td>34 (67%)</td>
<td>17 (33%)</td>
<td>24 (47%)</td>
<td>27 (53%)</td>
</tr>
<tr>
<td>Does trial coordinator feel responsible recruitment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes, very responsible</td>
<td>103 (53%)</td>
<td>91 (47%)</td>
<td>78 (40%)</td>
<td>116 (60%)</td>
</tr>
<tr>
<td>Not very responsible to responsible</td>
<td>67 (51%)</td>
<td>49 (42%)</td>
<td>61 (53%)</td>
<td>55 (47%)</td>
</tr>
<tr>
<td>Trial coordinator PhD</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>76 (51%)</td>
<td>74 (49%)</td>
<td>68 (45%)</td>
<td>82 (55%)</td>
</tr>
<tr>
<td>No</td>
<td>95 (59%)</td>
<td>67 (41%)</td>
<td>71 (44%)</td>
<td>91 (56%)</td>
</tr>
<tr>
<td>Requirements funder recruitment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>56 (62%)</td>
<td>35 (39%)</td>
<td>46 (51%)</td>
<td>45 (50%)</td>
</tr>
<tr>
<td>No</td>
<td>66 (50%)</td>
<td>66 (50%)</td>
<td>54 (41%)</td>
<td>78 (59%)</td>
</tr>
<tr>
<td>No funder</td>
<td>44 (54%)</td>
<td>38 (46%)</td>
<td>35 (42%)</td>
<td>47 (57%)</td>
</tr>
<tr>
<td>Was there a trial running with competing recruitment</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>45 (62%)</td>
<td>28 (38%)</td>
<td>37 (51%)</td>
<td>36 (49%)</td>
</tr>
<tr>
<td>No or unknown</td>
<td>136 (53%)</td>
<td>119 (47%)</td>
<td>110 (43%)</td>
<td>145 (57%)</td>
</tr>
</tbody>
</table>
PUBLICATION OF REGISTERED STUDIES

Forty-five percent of questionnaire respondents reported that (part of their) results had been published. For 1,031 of the 1,129 trials we searched for a published study report. We could identify a publication for 531 trials (52%), while for 471 (46%) we could not; for 29 trials (3%) we found an article on a very similar topic, by a very similar trial group, but we were uncertain whether they had reported the trial as registered, since a trial registration number was not reported, and there were unexplained differences in either in- and exclusion criteria or outcomes between the registry record and the study report.

DISCUSSION

In our analysis of unselected, registered trials we observed that enrolling the anticipated number of patients takes often longer than anticipated; even after a sometimes extended recruitment period the targeted sample size is often not reached. Contrary to our hopes, we could not identify a consistent pattern of predictors associated with recruitment success. The trialists in our survey suggested several potential solutions for improving trial recruitment, at multiple levels.

The studies in our analysis were sampled from a prospective trial registry (Netherlands Trial Register), and selection was not affected by publication of their results or based on the funder. As since 2005 registration of clinical trials is required for publication in most journals, we expect to have included a representative sample of trials.

A number of other issues in our analysis could invite discussion. Although the cohort of included trials represents an unselected sample, the response rate to our questionnaire was relatively low. This could have led to a preferential response of trialists that had encountered recruitment difficulties, since they may be more likely to respond to a questionnaire on these issues. One possible explanation for the low response rate could be that the trials had a planned date of completing recruitment between 2005 and 2010. Many of these trials had been registered about ten years ago, making it sometimes difficult to locate the trial coordinator, and the recruitment statistics of an already completed trial may have lost relevance for the investigator. We tried to partially remedy the low response rate by also looking at the recruitment statistics as published in the manuscript. Unfortunately, only about half of trials had published their study results two years or later after the planned completion date.
Our definition of successful recruitment – 80% or more of patients enrolled within the planned timeframe – could be considered arbitrary. We decided to integrate time into this outcome measure, as we see this as a crucial element when planning a trial. Others may consider recruitment successful if the targeted sample size was reached, even after an extension of the enrolment period. For this reason we also looked at reaching the full targeted sample at the actual stop date, regardless of its duration. Using that definition, only 53% of trials had recruited their targeted number of participants. We failed to find a consistent pattern of factors associated with recruitment success, or the lack thereof. A potential explanation may be the unselected nature of our sample, which we feel as a strong point, but also led to substantial heterogeneity between the trials in our analysis.

Slow trial recruitment has been frequently reported elsewhere in the literature but, so far, only a few studies have analysed associations between trial characteristics and recruitment performance in a large sample of studies.

In 1992 Easterbrook and colleagues looked at 720 research protocols approved by the REC in the United Kingdom, of which on 487 further information was obtained. They identified slow participant recruitment as the main reason given by the investigator for the study being abandoned or in abeyance (28% of 56 abandoned studies). In the STEPS study, Campbell and colleagues identified a sample of 114 studies, of which less than one third had recruited their original target within the time originally planned, and around one third had extensions. They found marginally statistically significant associations with being funded by the MRC, being a cancer trial and not having paid local trial coordinators, having a dedicated trial manager and an intervention being only available inside the trial. Yet confidence intervals were wide, reflecting the relatively small sample size.

Martin and colleagues looked at cardiovascular trials specifically, and found intensive trial related testing and an anticipated trial participation longer than six months to be associated with non-participation. Their analysis was on patient level – not on the trial level – and took only into account whether patients participated, without taking into account whether they were invited for participation or not.

Vale and colleagues studied 520 trials registered in a prospective UK cancer trial registry, UKCCCR NRCT. They report that, for the 333 trials that had completed recruitment 48% reached or exceeded the planned size, 19% recruited 75% or more of the planned numbers of patient, and 20% recruited less than 25% of the planned number of patients.
Toerien and colleagues performed a review of trials published in six major journals. They saw that 21% of 112 trials that reported a sample size calculation failed to achieve adequate numbers at randomization, and 48% at outcome assessment. As indicated by the authors, this will probably be an overestimation, given that all trials have been published in a high impact journal, and trials that have failed to achieve their sample size will probably less often be published in such a journal.

Although the help of local trial coordinators was perceived as one of the potentially most helpful strategies to improve recruitment mentioned by the respondents to our questionnaire, a negative association between paid local trial coordinators and successful recruitment was observed in the STEPS study. In our analyses we identified comparable negative associations with characteristics that could be hypothesized to be associated with better recruitment, like having a pilot or feasibility phase, presenting the study to recruiters or using trial pocket cards. It is likely that those interventions were only taken when recruitment difficulties were anticipated or observed. An alternative hypothesis is that having research staff will only improve recruitment if other necessary conditions are fulfilled, like support of key opinion leaders, and agreement on clinical relevance by patients and clinicians alike.

It is difficult not to agree on the large potential that exists for improving recruitment for trials at clinical sites. Unfortunately, we failed to identify a consistent pattern of trial characteristics associated with successful recruitment, though this is in line with previous, comparable attempts published in literature. This suggest that we should probably discontinue our search for simple, generally applicable predictors or ‘magic bullets’ that can guarantee success or failure in enrolling a sufficient number of patients.

A source of inspiration for alternative approaches could be the strategies that have been developed for improving implementation of evidence-based health care interventions in clinical practice. Several such strategies work with a series of plan-do-check-act cycles, in which one monitors whether a thoroughly designed recruitment plan, tailored to the trial, recruiting institute, patient population and expected obstacles, achieves the defined target. If not, one makes further modifications, as needed. McDonald and colleagues have proposed the use of a business model approach and marketing techniques for recruitment in clinical trials. The model seems promising, but more examples of its application in practice are needed. Expected obstacles could be derived from interviews
and focus groups with patients and health professionals, ran before the study is developed, and from pilot studies.\textsuperscript{24}

It is possible that recruitment is currently not considered as a priority among researchers, clinicians and patients. In that case, we should build strategies to convince and remind both health professionals and the general public of the value of research for practice. The efforts of health professionals to recruit patients for trials are often not financially rewarded in a similar way, treated differently from the act of treating patients, and introducing financial and legal triggers towards recruitment might be more effective.

Our study has confirmed that recruitment in clinical trials is slower than often expected, and perceived as difficult by trialists. Unfortunately our analysis of associations with recruitment success does not result in clear-cut recommendations for future trials. There does not seem to be a simple, effective package of measures that will lead to fail proof enrolment of the targeted number of patients in future trials. Most likely we will need a combination of multiple interventions, in a multifaceted approach, to recruit sufficient numbers of participants, and to provide clinical medicine with the evidence base that it requires to optimize health care effectiveness and quality.

ACKNOWLEDGEMENT

We would like to thank all trial coordinators for their response to the questionnaire.
REFERENCES


CHAPTER 4

WHY DO SOME CENTRES RECRUIT BETTER THAN OTHERS?

AN ANALYSIS OF RECRUITMENT RATES IN 17 RANDOMISED CLINICAL TRIALS IN OBSTETRICS AND GYNAECOLOGY

Katrien Oude Rengerink, Lotty Hooft, Noortje M. van den Boogaard, Birgit Y. van der Goes, Patrick M.M. Bossuyt, Ben Willem J. Mol
ABSTRACT

**Introduction** A commonly reported problem with the conduct of RCTs is that recruitment is usually slower than anticipated. We observe large differences between the numbers of patients recruited between different hospitals, even for similar trials. However, it is unclear which factors are associated with these recruitment differences.

**Design** We aimed to identify centre level factors associated with recruitment in hospitals participating in the Obstetrics and Gynaecology trials consortium in the Netherlands. We sent a web-based questionnaire to local researchers in centres that had recruited in total 14,808 patients for 17 multicentre trials. Our primary outcome was the summed weighted recruitment score of a hospital for all 17 studies, while the average weighted recruitment score per trial was a secondary outcome. The recruitment score was adjusted for the size of the catchment area. Using regression analysis we evaluated associations between recruitment scores and factors motivating centre decisions about participation in trials, the research orientation of the department, and the (perceived) logistic support by research personnel.

**Results** From 57 of 83 (69%) centres at least one questionnaire was returned. In univariable analysis, participating in trials because ‘expecting others to recruit in return for their studies’ was associated with higher recruitment scores, participating ‘because it is expected from us’ was associated with lower scores. Higher recruitment scores were seen in centres regularly initiating research and in academic medical centres. Weighted recruitment scores were higher in centres where research staff was available for more hours. On average coordinating centres recruited 4.2 times more patients than the median over all centres.

**Conclusion** Inclusion of patients in clinical trials is dependent on conviction of the value of recruitment for supporting practice and on the number of research staff. Our findings could be used to develop strategies aimed at improving recruitment rates in trials.
INTRODUCTION

Comparative effectiveness research is essential to inform evidence based health care decisions. The randomized controlled trial (RCT) is worldwide considered as the best instrument to evaluate the effectiveness of medical interventions. A commonly reported problem with the conduct of RCTs is that recruitment is usually slower than anticipated. In a cohort of 114 multicentre trials funded by the UK Medical Research Council and the UK Health Technology Assessment Programme between 1994 and 2002, less than one third recruited their targeted sample size within the time originally planned, and around one third needed more time.\(^1\) As a result, trials have less statistical power to detect potentially important differences between groups, which will make them less useful or even not used for guidance of decisions in clinical practice, if the targeted sample size is not achieved. Moreover, if recruitment has to be extended to reach the required sample size, the trial will cost more and take longer to complete.

As relevant adverse clinical outcomes in Obstetrics and Gynaecology are relatively scarce, large sample sizes are typically required to detect or refute relevant differences with limited statistical uncertainty. Therefore, in the Netherlands, up to 70 hospitals are working together in a consortium to recruit patients for a wide range of trials in women’s health. We observed large differences between hospitals in this consortium in the number of patients recruited, even for similar trials. As trial specific factors are the same in all centres, this variability may indicate that centre specific factors in conducting the trial play a role in recruitment for these trials. This phenomenon is not unique for our consortium.

Previous studies have shown that support from research staff positively influences recruitment, but other factors could also be influential.\(^2\)\(^-\)\(^5\) If barriers and facilitators for recruitment in centres could be identified, these could be linked to strategies to improve patient participation in trials within centres and eventually for selection of sites based on their recruitment performance. We collected centre information through a questionnaire and evaluated associations between centre characteristics and recruitment for trials organized by the Dutch Obstetrics and Gynaecology Trials Consortium.
METHODS

DESIGN
This study was part of the IMPACT study, of which the full protocol has been published elsewhere. We used a web-based questionnaire survey (available from the authors) to identify factors measurable at centre level potentially influencing recruitment in hospitals, both academic and non-academic, participating in the Obstetrics and Gynaecology Trials Consortium in the Netherlands.

PARTICIPANTS
A web-based questionnaire was sent to local research coordinators of all hospitals that had recruited for randomized trials run by the consortium, both academic and non-academic centres. To compare centre characteristics, we included all national multicentre randomized trials from our network, in which at least 5 centres had recruited patients, and that had finished or nearly finished recruitment when sending the questionnaire in September 2011. The selected trials were Allo, Amphia, Apostel2, Digitat, Hypitat, IUPC, Ppromexil, Probaat, STAN, Womb, ProTWIN, INeS, LifeSTYLE, Portret, Vusis1, Vusis2 and Trudil. All studies were performed in the field of Obstetrics and Gynaecology: 11 in obstetrics, 2 in reproductive medicine and 4 in urogynaecology. All studies had only recruited patients in hospitals in the Netherlands. Table 1 specifies the number of centres participating and the number of patients recruited for each trial. In total 14,808 patients had been included in these 17 trials, with a range of 60 to 5715. Eighty-three centres intended recruitment for at least one of these trials and were therefore eligible. Per trial between 6 and 52 centres had recruited for the trial. We sent the questionnaire to the local investigator or research coordinator in each of these 83 centres.

If required, we differentiated between obstetrics, reproductive medicine and urogynaecology, which implicated that in most cases questionnaires were sent to three subspecialties in a single centre and in some cases to one person responsible for all subspecialties. If the contact person did not feel he or she was the one best informed to answer questions, we asked to forward the questionnaire to a better informed colleague in the same centre. The questionnaire was not anonymous, which allowed us to send targeted reminders to non-responders. In case of non-response, two reminders were sent after the initial invitation.
QUESTIONNAIRE
The web-based questionnaire (available from the authors) was constructed based on literature and presented during a research meeting at the Academic Medical Centre for input from residents, research nurses and gynaecologists (in training). This questionnaire was then pilot tested by 3 research nurses, 4 gynaecologists and 4 residents, working in different subspecialties of Obstetrics and Gynaecology in a range of hospitals and adjusted based on the pilot experiences. The final questionnaire was then distributed to the above described participants with Survey Monkey [www.surveymonkey.com].

In the questionnaire we asked for three sets of factors potentially influencing recruitment in trials: A) Factors motivating centre decisions about participation in new trials; B) The research orientation of the department; and C) The (perceived) logistic support by research personnel.

Table 1: Subspecialty, sample size and number of centres of the trials included in the analysis

<table>
<thead>
<tr>
<th>Trial</th>
<th>Subspecialty</th>
<th>Sample size (N)</th>
<th>Centres (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allo\textsuperscript{d}</td>
<td>Obstetrics</td>
<td>210</td>
<td>11</td>
</tr>
<tr>
<td>Amphia\textsuperscript{g}</td>
<td>Obstetrics</td>
<td>665</td>
<td>44</td>
</tr>
<tr>
<td>Apostel2\textsuperscript{h}</td>
<td>Obstetrics</td>
<td>406</td>
<td>10</td>
</tr>
<tr>
<td>Digitat\textsuperscript{i}</td>
<td>Obstetrics</td>
<td>650</td>
<td>42</td>
</tr>
<tr>
<td>Hypitat\textsuperscript{j}</td>
<td>Obstetrics</td>
<td>756</td>
<td>35</td>
</tr>
<tr>
<td>IUPC\textsuperscript{k}</td>
<td>Obstetrics</td>
<td>1456</td>
<td>6</td>
</tr>
<tr>
<td>Promexil\textsuperscript{l}</td>
<td>Obstetrics</td>
<td>739</td>
<td>52</td>
</tr>
<tr>
<td>Probaat\textsuperscript{m}</td>
<td>Obstetrics</td>
<td>1176</td>
<td>21</td>
</tr>
<tr>
<td>STAN\textsuperscript{n}</td>
<td>Obstetrics</td>
<td>5715</td>
<td>10</td>
</tr>
<tr>
<td>Womb\textsuperscript{o}</td>
<td>Obstetrics</td>
<td>500</td>
<td>38</td>
</tr>
<tr>
<td>ProTWIN\textsuperscript{p}</td>
<td>Obstetrics</td>
<td>695</td>
<td>37</td>
</tr>
<tr>
<td>INeS\textsuperscript{q}</td>
<td>Subfertility</td>
<td>573</td>
<td>17</td>
</tr>
<tr>
<td>LifeSTYLE\textsuperscript{r}</td>
<td>Subfertility</td>
<td>475</td>
<td>23</td>
</tr>
<tr>
<td>Portret\textsuperscript{s}</td>
<td>Urogynaecology</td>
<td>463</td>
<td>23</td>
</tr>
<tr>
<td>Vusis 1\textsuperscript{t}</td>
<td>Urogynaecology</td>
<td>60</td>
<td>10</td>
</tr>
<tr>
<td>Vusis 2\textsuperscript{u}</td>
<td>Urogynaecology</td>
<td>128</td>
<td>20</td>
</tr>
<tr>
<td>Trudii\textsuperscript{v}</td>
<td>Urogynaecology</td>
<td>141</td>
<td>9</td>
</tr>
</tbody>
</table>
OUTCOME
Our primary outcome was the weighted recruitment (self-designed) score of a centre, summed over all studies. This score gives an overall score for recruitment within a centre, taking into account both centre size and the total number of trials that are recruiting concurrently at a centre. The summed weighted recruitment score was calculated as shown in Box 1. Higher scores indicate better recruitment. To adjust for the size of the catchment area, we standardized these numbers based on the number of clinical deliveries, fertility workups or urogynecological surgeries. The numbers used for standardization for centre size were provided to us by the local research coordinator in each centre. As an additional outcome measure we calculated an average recruitment score for each centre, by dividing the total weighted recruitment score by the number of trials performed in a centre. The analysis of these average recruitment scores was restricted to centres that had recruited for at least 3 trials.

We also looked at the potential for increasing recruitment, by looking at how the recruitment was in the coordinating centre compared to the other centres.

**Box 1: How is the weighted summed and average recruitment score calculated?**

- Take the number of patients recruited for the trial in the centre.
- To adjust for the size of the catchment area, standardize this number by dividing it by the number of deliveries (for trials in obstetrics), number of fertility workups (fertility) or number of urogynaecological operations (urogynaecology).
- The standardized measure is the centre and trial specific recruitment score.
- Calculate the trial-specific mean recruitment score over all centres participating in the trial.
- Calculate the deviation of the recruitment score from the trial-specific mean recruitment score, expressed in standard deviations.
- For the summed score: sum the mean recruitment score over all trials in which the centre has participated.
- For the average score: divide the sum over all trials by the number of trials the centre had recruited for.

STATISTICAL ANALYSIS
The analysis was performed at level of the centre. The first part of the questionnaire could be completed by more than one respondent, if in a centre there were different coordinators for the different subspecialties. If this was the case, only one answer per centre was selected. If answers were incongruent between respondents, the category most frequently mentioned was selected. If every one of the respondents in a centre had chosen a different response option, then the intermediate category was selected. Or, if it was a subspecialty specific question the answer of the person filling in the questionnaire for that subspecialty...
was selected if answers were incongruent.
When questionnaires were returned with missing response, we made no attempt at imputation of missing values. We calculated the number of centres selecting each response option in the questionnaire, and calculated percentages relative to the number of centres responding to the corresponding question.

We performed linear regression analysis, using the natural logarithm of the summed weighted recruitment score, as defined above, as the dependent variable. We have chosen the summed weighted recruitment score to be able to take into account the number of trials a site had recruited for, otherwise if a site recruited very well for only one trial will receive a high score, while a site that had recruited very well for nine trials but recruited average for one trial would receive a lower score. As independent variables we evaluated the factors potentially affecting recruitment mentioned earlier.

In a secondary analysis, we looked at the association between the factors and the average recruitment score recruited, in the subgroup of centres that had actually recruited patients for at least three trials.

RESULTS

At least one questionnaire was returned by 57 from the 83 (69%) centres for Obstetrics and Gynaecology. Data on recruitment rates of the 14,808 inclusions were available for all 17 trials and all centres. For data on recruitment we used all available data; for associations between factors in the questionnaire and recruitment data provided by the 57 responding centres were used.

Figure 1a shows the number of patients each centre had recruited for each trial, for 16/17 trials. For clarity and interpretability of the figure we excluded the STAN trial with over 5,000 inclusions. The median number of patients recruited was 62, with the 25th and 75th percentile being 5 and 139.

Figure 1b shows the recruited number of patients per hospital. As can be appreciated from this figure, the number varied widely both between hospitals and between studies. Of the 9403 inclusions shown, the 10 best recruiting centres had recruited 4968 (53%) patients, more than half of the total number of patients recruited. The 10 least recruiting centres did not recruit any patients, although they had expressed their intention to recruit for
these trials. Below we will discuss the strength of the associations separately for the three sets of factors defined earlier.

**Figure 1:** Overview of a) number of recruited patients for each of 16 trials; b) Number of recruited patients ranked for all centres for 16 trials

a. Number of recruited patient for each of the 16 trials*

* The x-axis shows the different studies, with on the y-axis the number of patients recruited for that study. Same colours represent the same centres. For clarity of the figure the STAN trial including over 5000 patients were excluded.

b. Number of recruited patients per hospital ranked for all centres for 16 trials, excluding STAN

* The y-axis shows the number of recruited patients. The x-axis shows the recruiting centers ranked in order of the number of recruited patients (y-axis).
A. Decisions regarding participation in trials

The decision to recruit patients for a trial was most often made during a staff meeting (47 centres; 79%), and/or by the person most specialized in that subspecialty (n=30; 53%). Twenty-one centres (38%) reported discussions between colleagues about participation; while 29 others (29%) indicated there was seldom such discussion. When asked to indicate how important the reasons to recruit for a study were, finding a solution for the clinical problem was selected as (very) important by all 56 responding centres. Other (very) important reasons for recruitment were authorship (n=28, 50%) and because they feel it is expected (‘It is expected from us’) (n=19; 34%). Three centres mentioned that participation depended on financial compensation, 9 centres reported it did not. In 41 centres the necessity of a financial compensation was dependent on other factors, mostly mentioned the actual costs. Six (11%) of the responding centres indicated they participate only if the research fits within the research priorities set by the hospital.

The number of studies running concurrently in a department was reported as limited in 24 centres; mostly to 2-3 trials, depending on work load. In the other 31 hospitals there was no such maximum. In 36 centres (63%) the number of trials a patient is asked for was limited, mostly to 1 or 2 studies.

Table 2 shows the results of the univariable analysis. Recruiting because ‘expecting others to recruit in return for their studies’ was significantly associated with higher recruitment scores. Recruiting because it was ‘expected to do so’ was significantly associated with lower scores. Authorship, financial compensation and being acquainted with the investigator were not influential.

When looking at the average recruitment score, restricting the analysis to the centers that had recruited for at least three trials, both ‘expecting others to recruit in return for their studies’ and ‘expected to do so’ were associated with the average recruitment scores.

B. Research orientation of the department

Twelve centers (21%) indicated never to initiate research themselves, but to participate in studies initiated by other hospitals. Regular discussions about research, as in journal clubs or research meetings, were organized weekly in 19 centers (34%), less than weekly in 33 (59%) and never in 4 (7%) of the centers. In 30 centers (53%) there was no time reserved for counseling during regular outpatient visits, but a special outpatient clinic run by the research staff was organized. In 1 (2%) there was always time reserved for counseling, in
6 (11%) at some occasions, and in 15 (26%) never. In 50 centers (89%) patient recruitment was discussed at the handover session: in 9 (25%) always and in 33 (63%) irregularly. Significantly higher recruitment scores were seen in centers regularly initiating research (Table 2). Higher scores were also seen in tertiary care centers. Having a research meeting was not found to be influential.

Table 2: Univariable association between factor and weighted recruitment score

<table>
<thead>
<tr>
<th>A</th>
<th>Decision regarding recruitment</th>
<th>Natural log of sum of z-scores per trial (N=57)</th>
<th>Average z-score if participating in ≥ 3 trials (N=36)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>B (95% CI) Ln transformed</td>
<td>B (95% CI) P-value</td>
</tr>
<tr>
<td>Dissension about participation sometimes/often vs no</td>
<td>0.20 (-0.03-0.44)</td>
<td>0.09</td>
<td>-0.15</td>
</tr>
<tr>
<td>Most important reason for participation</td>
<td>Finding a solution for a clinical problem</td>
<td>NA*</td>
<td>0.04</td>
</tr>
<tr>
<td>We get an inclusion fee</td>
<td>0.14 (-0.09-0.38)</td>
<td>0.22</td>
<td>0.16</td>
</tr>
<tr>
<td>Authorship of scientific article</td>
<td>0.15 (-0.08-0.38)</td>
<td>0.18</td>
<td>0.18</td>
</tr>
<tr>
<td>Recruitment is expected from us</td>
<td>-0.20 (-0.44-0.03)</td>
<td>0.09</td>
<td>-0.40</td>
</tr>
<tr>
<td>We know the (principal) investigator of the study</td>
<td>-0.06 (-0.31-0.19)</td>
<td>0.62</td>
<td>-0.33</td>
</tr>
<tr>
<td>We expect others to recruit in return</td>
<td>0.34 (0.11-0.58)</td>
<td>0.005</td>
<td>0.57</td>
</tr>
<tr>
<td>A maximum number studies in department concurrently</td>
<td>0.09 (-0.15-0.32)</td>
<td>0.463</td>
<td>0.04</td>
</tr>
<tr>
<td>A maximum number patient is asked for concurrently</td>
<td>-0.11 (-0.36-0.14)</td>
<td>0.375</td>
<td>-0.21</td>
</tr>
</tbody>
</table>

B Research orientation of the department

| Hospital type tertiary care versus other | 0.53 (0.23-0.84) | 0.001 | 0.68 | 0.01 |
| Regular research initiation of department yes/no | 0.30 (0.09-0.51) | 0.006 | 0.44 | 0.03 |
| Regular research meeting in department yes/no | 0.12 (-0.14-0.37) | 0.37 | 0.19 | 0.52 |
| Extra time for counselling during outpatient clinic yes/no | -0.01 (-0.35-0.33) | 0.95 | 0.32 | 0.36 |
| Outpatient clinic by research staff yes/no | 0.17 (-0.06-0.39) | 0.14 | 0.25 | 0.25 |
| Discuss recruitment during handover session always/not always | -0.05 (-0.23-0.14) | 0.62 | 0.08 | 0.90 |

C (Perceived) logistical support research staff

| Number of hours research staff appointed (hour) | 0.01 (0.002-0.02) | 0.03 | 0.01 | 0.08 |
| Research staff counsels patients in sub specialism Obstetrics (yes/no) | 0.07 | 0.24 | 0.08 | 0.42 |
| Fertility (yes/no) | -0.131 | 0.12 | -0.21 | 0.13 |
| Urogynaecology (yes/no) | -0.013 | 0.88 | -0.02 | 0.88 |

*The clinical problem was mentioned by all, therefore a measure of association could not be calculated.
Centers in which recruitment was always discussed during the handover session had the highest weighted recruitment score, followed by centers that discussed patient recruitment on an irregular basis during the handover session, but the difference was not significant.

When looking at the average recruitment scores, where the analysis was restricted to centers that had recruited for at least three trials, regularly initiating research was associated with higher average recruitment scores. Higher scores were also seen in academic medical centers.

C. (Perceived) logistical support by research personnel

In 47 centers (85%) research support staff was appointed for 1 up to 90 hours a week, mostly research nurses (N=24) and research midwives (N=15). Centers that did not have research support staff appointed all mentioned that it would be too costly, in view of the number of patients that was to be randomized.

Research nurses were reported to be working in 44 centers for studies in obstetrics, in 32 centers for subfertility trials, in 27 centers for trials in urogynaecology, and in three for oncology trials. Research staff was in 78% [28/36] available on fixed days only. Research staff was always available for counseling outside office hours in 2 centers (6%).

Weighted recruitment scores were significantly higher in centers when more hours of research staff were available (p=0.03). When adjusted for type of hospital (academic versus non-academic), the difference disappeared. Whether or not research staff also counseled patients did not influence the weighted recruitment score in a statistically significant way. When looking at the average recruitment score, restricting the analysis to the centers that had recruited for at least three trials, the number of hours a research nurse was appointed was also significantly associated with higher recruitment scores. When adjusted for the number of studies performed in a center this effect was no longer significant (p-value of 0.08).

In Appendix 1 the number of patients recruited in the coordinating center is compared to the median number of patients recruited. On average the coordinating center recruits 4.2 times more patients than the median number of patients (calculated over all participating centers), it recruits on average 0.8 time more than the best recruiting center, and 42 times more than the least recruiting center. These numbers were not corrected for any other potentially influential factors, such as the length of the recruitment period (initiating centers typically start earlier).
DISCUSSION

In this study, we found that research oriented departments in which physicians were convinced of the value of recruitment for supporting practice recruit more patients in clinical trials, as do centres with more research staff appointed. Recruitment in trials because their centre primarily feels they were expected to do so was associated with lower recruitment scores. On average the coordinating center recruits 4.2 times more patients than the median number of patients.

Strong points of our study are that we included a variety of trials and a large series of centres, in which in total 14,808 patients were recruited. The high number of centres made it possible to compare recruitment rates between centres for specific trials. We would also like to mention a few potential limitations. We analysed the questionnaires at the centre level, not taking into account the perception of all individuals in that centre. We selected this strategy, as we hoped to get a response from the best informed gynaecologists in a centre, without bothering all of the colleagues.

For most studies the actual number of patients eligible for a trial was unknown. To adjust for the size of the catchment area, we standardized these numbers. These adjustments may not reflect smaller differences in eligibility, due to variability in the prevalence of inclusion and exclusion criteria. We did not take into account the time period a centre was recruiting for a trial. We assumed that all trials could have started the process of getting ethical approval at the same time. We are aware that the coordinating centre often starts the study in its own centre early, to test whether all procedures are clear and running smoothly.

Haidich and colleagues evaluated recruitment in early and late-starter sites in a cohort of 14 randomized trials conducted by the Adult AIDS Clinical Trials Group and found that sites that started recruitment within 5 months from the time the first patient entered the trial were eventually responsible for over 90% of the total enrolment in 11 of the 14 trials. They concluded that the late-starter sites were unlikely to make important contributions to eventual trial enrolment in large clinical trials conducted by groups with a fixed number of sites. Given the common stop date in all trials, this indicates that adjustment for the length of the recruitment period could even unintentionally give a too optimistic picture of the worst recruiters.

We found that centres with research staff appointed recruited better. Studies in different
clinical areas have also shown that support from research staff positively influences recruitment. However, Campbell et al. observed lower recruitment rates in trials with paid local recruitment staff. Tarnow-Mordi et al. pointed out that government funded clinical research networks, including local site research nurses or coordinators to support clinical trials, could enhance recruitment in sites with an adequate volume of patients. Such networks have already been established for cancer trials in Australia and for pediatric trials in the United Kingdom. Although it seems very promising, it’s cost-effectiveness has not been proven yet. We found that research oriented departments in which physicians were convinced of the value of recruitment for supporting practice recruit more patients. De Wit and colleagues found that in primary care research successful patient recruitment is more affected by the motivation level of the research group than by financial incentives, the research topic, or research experience.

Van Kuyvenhoven and colleagues observed that 50% [4/8] of the general practitioners who participated because they knew the researchers did not recruit any patients, compared with 16% [9/58] that participated for other reasons. Wilson and colleagues found that general practitioners that recruited for trials were more interested in knowing more about research than those who did not recruit. Ziebland et al. concluded that it does matter if clinicians do not understand the rationale for the trial or if they view the results as ultimately irrelevant to their practice. Financial incentives as a reason for recruitment were not significantly associated with recruitment rates. A review by Bryant and Powell concluded there is very limited, inconclusive evidence about financial incentives to improve recruitment.

These results also invite a discussion on the optimal number of centers that should be asked to recruit for a trial. Recruiting in more sites might be faster, but might be a less efficient use of resources. In our results, the 10 best recruiting centres recruited more than half of the total, while the 10 least recruiting centres hardly contributed. Therefore, it might be more efficient to select a lower number of centres for recruitment. However, recruiting for a trial can also give a slight increase in implementation of the results of the trial.
Based on our results, we see a large potential for improving recruitment rates at clinical sites. Logistical infrastructure is a prerequisite for executing trials, but not a sufficient condition for trial recruitment. Sites should be convinced of the value of research for practice and encouraged to consider appointing a research nurse in order to increase recruitment rates. Clinical trial recruitment should be embedded in routine patient care, as trials are the foundation for building an efficient, affordable and high quality health care system.

ACKNOWLEDGEMENTS
We would like to thank all respondents to the questionnaire, all study coordinators for using their recruitment rates and the residents, gynaecologists and research staff who pilot tested the questionnaire.
REFERENCES


Appendix 1: Number of recruited patients in the coordinating centre compared to the median number of patients recruited.

<table>
<thead>
<tr>
<th>Study</th>
<th>Nr of patients recruited in coordinating centre</th>
<th>Median recruited number over all centres (min-max)</th>
<th>Number of times coordinating centre recruits better than median over all centres (better than best centre-better than worst centre)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allo</td>
<td>53</td>
<td>19 (1-53)</td>
<td>2.8 (1.0-53.0)</td>
</tr>
<tr>
<td>Amphia</td>
<td>36</td>
<td>10 (1-51)</td>
<td>3.6 (0.7-36)</td>
</tr>
<tr>
<td>Apostel2</td>
<td>30</td>
<td>30 (5-87)</td>
<td>1.0 (0.34-6.0)</td>
</tr>
<tr>
<td>Digitat</td>
<td>16</td>
<td>9 (1-74)</td>
<td>1.8 (0.22-16.0)</td>
</tr>
<tr>
<td>Hypitat</td>
<td>52</td>
<td>17 (2-69)</td>
<td>3.1 (0.75-26)</td>
</tr>
<tr>
<td>IUPC</td>
<td>412</td>
<td>173 (73-552)</td>
<td>2.4 (0.75-5.6)</td>
</tr>
<tr>
<td>Ppromexil</td>
<td>31</td>
<td>10 (0-35)</td>
<td>3.1 (0.89-31)</td>
</tr>
<tr>
<td>Probaat</td>
<td>213</td>
<td>43 (10-213)</td>
<td>5.0 (1.0-21.0)</td>
</tr>
<tr>
<td>STAN</td>
<td>1350</td>
<td>435 (13-1471)</td>
<td>3.1 (0.92-104)</td>
</tr>
<tr>
<td>WOMB</td>
<td>68</td>
<td>6 (0-76)</td>
<td>11.3 (0.89-68)</td>
</tr>
<tr>
<td>ProTwin</td>
<td>87</td>
<td>16 (1-87)</td>
<td>5.4 (1-87)</td>
</tr>
<tr>
<td>INeS</td>
<td>88</td>
<td>29 (2-88)</td>
<td>3.0 (1-44)</td>
</tr>
<tr>
<td>LifeStyle</td>
<td>39</td>
<td>18 (2-62)</td>
<td>2.2 (0.7-19.5)</td>
</tr>
<tr>
<td>Portret</td>
<td>136</td>
<td>13 (2-136)</td>
<td>10.5 (1-68)</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td>****</td>
<td><strong>4.2 (0.80-42)</strong></td>
<td><strong>------------------------------------------------------------------------------------------------</strong></td>
</tr>
</tbody>
</table>

Appendix 1: Number of recruited patients in the coordinating centre compared to the median number of patients recruited.
CHAPTER 5

PREGNANT WOMEN’S CONCERNS WHEN INVITED TO A RANDOMIZED TRIAL

Katrien Oude Rengerink, Sabine L.M. Logtenberg, Lotty Hooft, Patrick M.M. Bossuyt, Ben Willem J. Mol
ABSTRACT

Introduction Although pregnant women were until the 1990’s excluded from clinical trials, the Food and Drug Administration nowadays allows - and even encourages - responsible inclusion of pregnant women in trials with adequate safety monitoring. Still, randomized trials in pregnant women may face specific enrolment challenges.

Methods We performed a qualitative case control study. Women who had been recently invited to one of eight clinical trials during pregnancy or shortly after giving birth were invited for a face-to-face interview, to identify their motives for participation. We selected both participants and non-participants, in a 1:1 ratio. We started the interview in an open fashion, asking for the women’s main motive for participation or non-participation. When no new information emerged we continued with a semi-structured interview, guided by a topics list. Transcripts of the interviews were analysed using a constant-comparative approach. Two researchers identified barriers and facilitators for participation, conjoined into main themes.

Results Of 28 participants invited to this study, 21 consented to be interviewed. Contribution to scientific research was for 5 of 12 participants the main motive for participation in the trial, while 5 mentioned to have participated because the intervention was not available outside the trial. Key motives for non-participation (n=9) were a negative association or dislike of the intervention, either because it might do harm (n=6) or for practical reasons (n=3). Combining the open and topic-list guided interviews led to 47 sub-codes, which we conjoined into seven main themes: external influence, research and healthcare, perception own situation, study design, intervention, information and counselling, and uncertainty.

Conclusion We identified seven main themes that seem to influence their decision about participation. We noted that uncertainty about scientific research and/or the intervention was reported to be of considerable importance. New studies should look into methods to further reduce the feeling of uncertainty around trial participation decisions by pregnant women.
INTRODUCTION

Up to the 1990s, pregnant women were often excluded from clinical trials, for their own protection.\(^1\) However, pregnancy does in general not prevent or cure a woman from (acquiring) a disease. The efforts to protect the foetus from research-related risks, by excluding pregnant women from research, places worldwide women and their foetuses at risk from unstudied interventions.\(^2,3\)

In the United States about 2 in 3 of pregnant women are given a (off-label) prescription medication during pregnancy, often based on limited evidence on safety or effectiveness.\(^2\) Effectiveness research is needed to inform evidence-based healthcare decisions, and results of studies in non-pregnant women may not always apply to pregnant women. The Food and Drug Administration nowadays allows – and even encourages – responsible inclusion of pregnant women in drug trials with adequate safety monitoring.\(^4\)

Randomized trials in pregnant women may still face specific enrolment challenges. Enrolment problems are not limited to studies in pregnant women, but studies including pregnant women are unique, since two patients are involved: the mother and her unborn foetus. A woman may refuse treatment for herself if she feels this could harm her baby, or she may feel bound to accept interventions that might benefit the foetus. Additionally, the father’s feelings may also influence decision-making about participation in a trial.\(^5\)

Tooher and colleagues performed a narrative review on factors influencing recruitment for maternal and perinatal trials.\(^6\) They identified four participant factors that influence recruitment: understanding of risk, recruitment process and procedures, participants understanding of the research process and methodological issues, and patient characteristics. Their conclusions were based on a limited number of studies on maternal trials, and several of these were performed because recruitment was problematic. It is uncertain to what extent these results also apply to other studies. Moreover, trial participation considerations may have changed over time and could differ between countries. We performed a qualitative study to identify main barriers and motivators for enrolment in obstetrical trials in the Netherlands.
Chapter 5 | Pregnant women’s concerns when invited to a randomized trial

METHODS

DESIGN
We performed a qualitative case-control study. Women recently invited to clinical trials during or shortly after pregnancy were invited for a face-to-face interview about their main motives to accept or decline the invitation to participate in a RCT. This study is part of the IMPACT study, in which enrolment of patients in trials is studied at different levels. Our study did not require formal approval of an ethics committee or internal review board, according to Dutch law, as confirmed by the ethics committee of the Academic Medical Centre and the Onze Lieve Vrouwe Gasthuis.

SELECTION OF TRIALS AND INVITATION OF INTERVIEWEES
We identified women invited less than 3 months ago to enrol in a clinical trial in obstetrics and sampled, in a 1:1 ratio, stratifying for whether they had accepted or had declined enrolment. These women were selected from eight multicentre studies that had been actively recruiting patients between February and June 2010, all running in the Consortium for Women’s health and Reproductivity studies: Allo, Apostel I, Apostel II, Chips, WOMB, Ppromexil, Hypitat2, and ProTwin trial. A short description of these trials is shown in Table 1; more information about these studies can be found at www.studies-obsgyn.nl.

Women were eligible if they were still pregnant or their baby was born alive and they were able to speak the Dutch language well enough to participate in the interview, without an interpreter. For practical reasons, only patients from three different geographical areas were contacted: Amsterdam (Academic Medical Centre; St Lucas Andreas Hospital; Onze Lieve Vrouwe Gasthuis), Enschede (Medisch Spectrum Twente) and Veldhoven (Maxima Medical Center). We started our invitations with the women most recently invited to the trials. One of the interviewers worked as a clinical midwife, and as such was also responsible for recruitment of patients in the Onze Lieve Vrouwe Gasthuis. Excluded were women who had a professional relation with the interviewer. Interviewees were initially contacted via a letter, sent on behalf of both their gynaecologist or the local trial coordinator, and the interviewers. After about a week the patients were contacted by phone, and invited for an interview.

THE INTERVIEW
The interview was run face-to-face, unless the respondent explicitly requested a telephone
interview, or when the travel time to visit the patient was 2 hours or more. The interview took place at the patient’s home, or in the hospital, whichever was preferred by the interviewee.

We started the interview in an open fashion, by asking the women for their main reason for participation, or non-participation, in the trial. When this open part did not produce any new information, we continued with a semi-structured interview. This section was guided by a topic list, to cover all aspects that might have contributed to the decision making process. This topic list, available in Appendix 1 (Dutch), targeted potential barriers and facilitators for (non)participation. It was developed based on a literature review and with input from experienced gynaecologists and midwives. The topic list included factors related to personal benefit, altruism, knowledge and information about the trial and the trial process, distrust, attitude, organisation aspects and influence of the social environment. If new topics emerged during the interview, they were added to the topic list.16,17

The total number of interviews was not set, but depended on data saturation.17 We estimated that an interview with 5 to 10 women in both groups would be needed to reach saturation. We planned to perform two additional interviews when data saturation was reached. We also collected maternal ethnicity, age, parity, educational level, height and weight before pregnancy.

ANALYSIS
All interviews were recorded with a voice-recorder and transcribed; explicit non-verbal communication was noted. The transcribed interview was sent to the interviewee, and we asked her to confirm its correctness and completeness (member-check). Transcripts of the interviews were analysed using Microsoft Excel. The aim of the analysis was to conceptualize the content of the interviews in main themes. Analysis was performed according to the taxonomy of Strauss and Corbin (‘create theory out of data’), where one starts with line-by-line open coding of all relevant phrases of barriers or motivators for participation (open coding), using a constant comparison method: newly gathered data are continually compared with previously collected data and their coding in order to refine the development of theoretical categories.

After this open coding, the codes were grouped into subcategories (axial coding), and conjoined into themes (selective coding).17 All transcripts were reread and recoded, using the improved coding structure to ensure no codes were missing. If a fragment fitted more
subcategories, the fragment was placed into all relevant categories.

Two researchers (SL and KOR) independently marked barriers and facilitators for participation for the first seven interviews. Thereafter one researcher marked barriers and facilitators, checked by a second researcher, and dissolved by consensus if needed. These phrases of barriers and facilitators were classified into categories and conjoined into themes.

RESULTS

INTERVIEWEES
We sent 28 women an invitation by mail. When we called the women thereafter, to answer any remaining questions and be informed about their decision on participation, 4 women could not be reached after four or more attempts. Two women declined the invitation for an interview (reason not noted); one woman who initially consented to an interview was admitted to the hospital for emergency care and her interview was cancelled.
In total 21 interviews were performed; 12 with trial participants and 9 with women who declined participation. Of these, 17 were face-to-face interviews and 4 interviews were by phone (in 3 cases because this was requested by the interviewee and in one case because of the distance).

The interview took on average about half an hour. After transcription of the recorded interview, 20 of 21 interviewees approved its content, 1 woman did not respond. Characteristics of the interviewees are shown in Table 2.

Although our inclusion criteria were designed to select only patients invited to enroll in a trial no longer than three months ago because of potential recall bias, five women were invited for this study more than three months ago after their RCT invitation. This was mostly due to the registration of non-participants in a trial, which was not always complete. All respondents stated they remembered the situation to be discussed in the interview very well, which we could confirm during the interviews.
**Table 1: Overview of trials from which patients were selected for an interview.**

<table>
<thead>
<tr>
<th>Trial acronym*</th>
<th>Research question</th>
<th>Treatment arms</th>
<th>Eligible women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allo*</td>
<td>Does antenatal allopurinol administration reduce hypoxic-ischaemic encephalopathy in neonates exposed to intra-uterine asphyxia?</td>
<td>Allopurinol or placebo, antenatal administered to the mother</td>
<td>Women at term in whom the foetus is suspected of intra-uterine asphyxia</td>
</tr>
<tr>
<td>Apostel I†</td>
<td>Is testing for fibronectin a cost-effective strategy that prevents unnecessary treatment in women with threatened preterm labour?</td>
<td>Tocolytics (nifedipine) or placebo</td>
<td>Patients with symptoms of preterm labour, and a negative fibronectin test and a cervical length between 10-30 mm</td>
</tr>
<tr>
<td>Apostel II†</td>
<td>Does sustained tocolysis in women with threatened preterm labour reduce neonatal morbidity?</td>
<td>Nifedipine or placebo for 12 days</td>
<td>Women between 24 to 31** weeks pregnant who have been treated with tocolysis and steroids for preterm birth for 48 hours</td>
</tr>
<tr>
<td>CHIPS**</td>
<td>Is there a difference on pregnancy loss or NICU admission between less tight and tight control of blood pressure in women with non-severe non-proteinuric pre-existing hypertension or gestational hypertension remote from term?</td>
<td>‘less tight’ dBP control or ‘tight’ dBP control</td>
<td>Women with non-severe non-proteinuric pre-existing hypertension or gestational hypertension remote from term</td>
</tr>
<tr>
<td>Hypitat II†</td>
<td>What is the effectiveness and efficiency of induction of labour in women with pregnancy induced hypertension or mild preeclampsia with a gestational age of 34-37 weeks of pregnancy, as compared to expectant management under regular monitoring?</td>
<td>Induction of labor or expectant management under regular monitoring</td>
<td>Women with pregnancy induced hypertension or mild preeclampsia with a gestational age of 34 - 37 weeks of gestation</td>
</tr>
<tr>
<td>Ppromexil††</td>
<td>What is the effectiveness and cost-effectiveness of induction of labor after PPROM between 34 and 37 weeks gestation compared to expectant monitoring?</td>
<td>Induction of labor or expectant monitoring</td>
<td>Pregnant women with preterm premature rupture of membranes between 34 + 0/7 weeks to 37 weeks of gestation</td>
</tr>
<tr>
<td>ProTWIN††</td>
<td>Is prophylactic use of a cervical pessary effective in the prevention of preterm delivery and the neonatal mortality and morbidity resulting from preterm delivery in multiple pregnancy?</td>
<td>Pessary or no treatment.</td>
<td>All women presenting with a multiple pregnancy between 12-20 weeks of gestation</td>
</tr>
<tr>
<td>WOMB‡‡</td>
<td>What is the effect of RBC transfusion on health related quality of life?</td>
<td>RBC transfusion or no intervention</td>
<td>Women with PPH or a decrease in Hb, 12 to 24 hours after delivery or caesarean section</td>
</tr>
</tbody>
</table>

*More information about these studies can be found at: www.studies-obsgyn.nl*
Table 2: Characteristics of the women included

<table>
<thead>
<tr>
<th>Code</th>
<th>Ethnicity</th>
<th>Level of education</th>
<th>Age</th>
<th>Study</th>
<th>Parity*</th>
<th>Place</th>
<th>Hospital</th>
</tr>
</thead>
<tbody>
<tr>
<td>J</td>
<td>Dutch</td>
<td>Intermediate/low</td>
<td>30</td>
<td>Allo</td>
<td>0</td>
<td>Veldhoven</td>
<td>MMC</td>
</tr>
<tr>
<td>L</td>
<td>Dutch</td>
<td>Higher education</td>
<td>39</td>
<td>Allo</td>
<td>1</td>
<td>Veldhoven</td>
<td>MMC</td>
</tr>
<tr>
<td>M</td>
<td>Dutch</td>
<td>Higher education</td>
<td>34</td>
<td>Allo</td>
<td>1</td>
<td>Veldhoven</td>
<td>MMC</td>
</tr>
<tr>
<td>N</td>
<td>Dutch</td>
<td>Higher education</td>
<td>29</td>
<td>Apostel I</td>
<td>2</td>
<td>Veldhoven</td>
<td>MMC</td>
</tr>
<tr>
<td>P</td>
<td>Dutch</td>
<td>Higher education</td>
<td>28</td>
<td>Ppromexil</td>
<td>0</td>
<td>Enschede</td>
<td>MST</td>
</tr>
<tr>
<td>Q</td>
<td>Dutch</td>
<td>Intermediate/low</td>
<td>26</td>
<td>Ppromexil</td>
<td>0</td>
<td>Enschede</td>
<td>MST</td>
</tr>
<tr>
<td>T</td>
<td>Dutch</td>
<td>Higher education</td>
<td>32</td>
<td>Hypitat II</td>
<td>1</td>
<td>Amsterdam</td>
<td>OLVG</td>
</tr>
<tr>
<td>U</td>
<td>Dutch</td>
<td>Higher education</td>
<td>39</td>
<td>ProTWIN</td>
<td>0</td>
<td>Amsterdam</td>
<td>OLVG</td>
</tr>
<tr>
<td>V</td>
<td>Surinam</td>
<td>Higher education</td>
<td>38</td>
<td>Hypitat II</td>
<td>2</td>
<td>Amsterdam</td>
<td>OLVG</td>
</tr>
<tr>
<td>W</td>
<td>Dutch</td>
<td>Higher education</td>
<td>36</td>
<td>Apostel II</td>
<td>2</td>
<td>Amsterdam</td>
<td>OLVG</td>
</tr>
<tr>
<td>Y</td>
<td>Surinam</td>
<td>Higher education</td>
<td>37</td>
<td>CHIPS</td>
<td>1</td>
<td>Amsterdam</td>
<td>OLVG</td>
</tr>
<tr>
<td>Z</td>
<td>Belgian</td>
<td>Higher education</td>
<td>29</td>
<td>WOMB</td>
<td>2</td>
<td>Amsterdam</td>
<td>SLAZ</td>
</tr>
<tr>
<td>A</td>
<td>Dutch</td>
<td>Higher education</td>
<td>35</td>
<td>WOMB</td>
<td>1</td>
<td>Amsterdam</td>
<td>SLAZ</td>
</tr>
<tr>
<td>B</td>
<td>Dutch</td>
<td>Unknown</td>
<td>27</td>
<td>ProTWIN</td>
<td>0</td>
<td>Amsterdam</td>
<td>OLVG</td>
</tr>
<tr>
<td>C</td>
<td>Dutch</td>
<td>Higher education</td>
<td>34</td>
<td>ProTWIN</td>
<td>1</td>
<td>Amsterdam</td>
<td>SLAZ</td>
</tr>
<tr>
<td>D</td>
<td>Moroccan</td>
<td>Higher education</td>
<td>33</td>
<td>ProTWIN</td>
<td>1</td>
<td>Amsterdam</td>
<td>SLAZ</td>
</tr>
<tr>
<td>E</td>
<td>Surinam</td>
<td>Higher education</td>
<td>39</td>
<td>Hypitat II</td>
<td>1</td>
<td>Amsterdam</td>
<td>AMC</td>
</tr>
<tr>
<td>G</td>
<td>Dutch</td>
<td>Higher education</td>
<td>27</td>
<td>Apostel II</td>
<td>0</td>
<td>Amsterdam</td>
<td>OLVG</td>
</tr>
<tr>
<td>I</td>
<td>Dutch</td>
<td>Higher education</td>
<td>22</td>
<td>ProTWIN</td>
<td>0</td>
<td>Enschede</td>
<td>MST</td>
</tr>
<tr>
<td>K</td>
<td>Dutch</td>
<td>Intermediate/low</td>
<td>22</td>
<td>Ppromexil</td>
<td>1</td>
<td>Veldhoven</td>
<td>MMC</td>
</tr>
<tr>
<td>O</td>
<td>Dutch</td>
<td>Intermediate/low</td>
<td>23</td>
<td>Hypitat II</td>
<td>0</td>
<td>Veldhoven</td>
<td>MMC</td>
</tr>
</tbody>
</table>

*Parity was registered at the time of the interview.

MAIN MOTIVE FOR TRIAL PARTICIPATION OR NON-PARTICIPATION

Contribution to scientific research was for 5 of the 12 participants the main motive for participation in the trial, as responded to the first open question:

Mw J “I think research projects are actually never bad, and this is not a type of research where they do real experiments, so it is always good to learn from this someone else”

Mw L “In our first pregnancy our daughter was in fetal distress, and then it became a caesarean section. Then maybe this had been an option as well, because it is associated with fetal distress, and then administer this. And for my husband it was actually by asking further: does it have disadvantages for the child? No? Then we participate, because the study is also necessary. .... Also my medical background, I also worked on the labor wards. You are working in medicine, so you are open for innovation and technique.”
Mw T “There are two issues: in my first pregnancy I had preeclampsia, so I was very well aware what the consequences would be for me, and then also for the child. ... For me, I supported the aim of the study, so deliver the baby from 34 weeks, because the maternal and fetal risks do not outweigh, so to say. Second, I am in an academic hospital, already for years, also for other treatments, and I believe very much in academic. I believe in development and trying things. Well, and research is part of that, because if you do not do any studies, you can never do something new.”

Mw U “Well at first I was invited to participate in a study about the pessary, to participate in a twin study. Well, I thought, seems good, I have a little twin sister and I got pregnant with ICSI, so for me there were also people who participated in such a study, because of which I am pregnant now.”

Mw Z “Well, first, It was very much applicable, and it was the choice between taking blood or iron, and otherwise it would anyway be iron, so I thought let’s see what happens. And I was in the blood group, in retrospect I was very happy with it. And, what I just said, I do studies myself often, yes, than you better see the importance of it, that you need to recruit people, so eh, that is the only motivation.”

Five participants mentioned to have participated because the intervention was not available outside the trial:

Mw M “Most of influence is of course that the consequences of shortage of oxygen are pretty fierce, if you could reduce that somehow by taking a certain substance, I would choose that. Yes, yes, well and because the substance was already in use for other purposes, yes, it is not fully tested for shortage of oxygen, that it should not be harmful, than its your assumption that it therefore would only have advantages. And then I think something like, I want to participate in that study.”

Mw P “Okay, well, it was mainly because the fact that there was a chance that my labor would be induced, otherwise I would have to wait till 37 weeks anyway .... Of course also the reason to contribute to research, that was also a good objective, but that was not the most important thing. 'I thought: I want to go for immediate induction’. I could not imagine that I had to stay in the hospital for five days, because I was not allowed to do anything, and I thought like: “bring it on...”"
Mw Q: “At a first glance: I do not imagine myself lying here for another 5 weeks, and pretty soon thereafter the idea that you are already open from down there, and a risk of infection for yourself and for the baby, and yes, in Enschede the doctors also said: the baby was viable enough, so that was a reason to participate for us.”

Mw V: “That if the baby would be born, that my high blood pressure would be gone, that’s what I thought, that was about it. But on the other side, I was a bit scared, shall I get it earlier, that was at 36 weeks, so it was a bit the consideration what would be best. Then she explained me, the earlier the child would be out, the better it could be for mother and child, so that was actually the reason that I said: I participate.”

Mw W: “That was because I hoped it would be better for the child, although I had an uneasy feeling all the time. That was because nobody could tell you what the potential negative consequences were, yes, I had an uneasy feeling all the time.”

One women thought an extra test could only be positive, a kind of ‘there is no harm in trying’:

Mw N “And I had something like, in my case it can only be positive, because I mean, the test would indicate whether the chance was very high that you would deliver very soon, or that it could take a while. So, I really felt like, I felt that I ran little risk, because if the test would show that you would fall into the test group, than you would get either a placebo or tocolytics”.

For one women the reason was not very clear, she probably meant to be better informed about her medical condition:

Mw Y “Than you know how and what”.

Key motives for the 9 non-participants were a negative association or dislike of one of the interventions, either because it might do harm (n=6):

Mw C “Well there were multiple reasons. When your colleague started about it, when I had an appointment about it, I thought ‘Oh my God, no, not a pessary. Because I had a friend who was admitted to the hospital because the pessary [not in pregnancy] had caused many bleedings, so that’s what I told her, that was a life threatening situation, so I had a feeling like, if I think now about pregnancy and a pessary, I do not get very happy.”
Mw D: “For me it was pretty clear actually, when I was here, I thought something like ‘let mother nature just do the work. I am pretty religious (Muslim), but there might be a reason why the children are born early, I believe in God you know, I have something like: destiny determines, actually, if the children want to be born early, then that is the case, if not it is not. That was my consideration. I was also scared, if I participate and something happens with me, a bit or a lot blood loss, or something with the babies.”

Mw E: “She (baby) was 4 weeks too early and the blood pressure kept rising, they did not get it down. I was lying there for a month, and I had enough of it, you want something to change. Then they asked me: do you want to participate in the study? At a certain moment the doctors said, we don’t know it anymore. Then I thought: if they don’t know, who does, I had to choose myself. ….. And then I thought, actually: ‘I can better prolong it for a while, look how long it will take. Because if I had decided to participle, will you be induced or not, that is also an uncertainty. So, then the disappointment is still big. Then I decided not to do it, let’s have a look how long we can prolong it.”

Mw G: “I had a very tough pregnancy, with a lot of bleeding, and actually the nine months were completely uncertain. … I got lung maturation injections and I got tocolysis and then I did not feel my baby any longer. … When they asked if I would receive more tocolysis, I associated it with that, and I wanted to feel my baby again as soon as possible, to get the certainty back a bit, that everything was alright. So that was for me the most important reason.”

Mw I: “At first, I tended to participate, because in my environment many people said: imagine it works, your babies will stay in longer. But I had the feeling that the pregnancy goes very well, that it all, yes I react quite strong on things, jewels or a piercing or something, then I think, if something is brought into my body, maybe it might react strangely. If nothing is wrong and you do that, that’s a bit scary. And that you could not choose which group you will get in, that pretty logical in a study, but that’s why I finally decided not to.”

Mw K: “Well, it’s not without a reason that they tell you that from 37 weeks you are officially allowed to deliver, so, yes, I thought it was a risk to be induced at 34 weeks. Because the doctors do not say without a reason, from 37 weeks doctors will automatically induce you, and they are doing a study, and I did not want to be a guinea pig. If then something goes wrong...”
Chapter 5 | Pregnant women’s concerns when invited to a randomized trial

Or for practical reasons associated with the intervention (n=3):

Mw A “I wanted to participate if I could choose for iron tablets, but that choice was not there. You have to participate blindly and then determines I don’t know who, I don’t know how that works, but then determines someone else for you which of the two you are going to do. What is also complicated, I did not want a blood transfusion. I was lying there on a drip and I had a catheter and then I thought that with iron tablets I could go home and otherwise I would have stayed somewhat longer.”

Mw B “I was in the AMC, that is an academic hospital, and there they do a lot of research. I had to come extra, it was for example about a pessary for prevention of preterm birth, I had to come extra to measure it and for an ultrasound. I did seriously consider it, but the extra visits, for example if I would have pain or if it would not fit well. And, I just heard I was pregnant of a twin with 1 amnion, that is a very exceptional situation, a lot of information is coming to you”.

Mw O [unplanned pregnancy]: “.....Yes, and everything went very fast, then I had really something like, well I do not have to induced tomorrow. That... the chance was 50% and I did not want it. No, that’s all too soon. Because you are.. no... you are after three weeks that I was attending the hospital I was admitted. I had never been admitted in a hospital. I was homesick, yes. But, I did not have something like, get him tomorrow. That was too soon. I could not process that”.

Women also mentioned it played a role that it was their first pregnancy, or she was already in an exceptional situation given a mono-amniotic twin pregnancy.

THEMES IDENTIFIED AS RELATED TO THE DECISION ON TRIAL PARTICIPATION
During the phase of open coding 47 subcategories were identified, based on phrases relating to barriers and facilitators. These subcategories were aggregated into 13 main categories, and further classified into seven main themes: (1) external influence, (2) research and healthcare, (3) perception own situation, (4) study design, (5) intervention, (6) information and counseling, and (7) uncertainty. These main themes, with corresponding sub codes, are summarized in Table 3. Each of the themes is discussed separately below.
A. External influence
Women indicated that they discussed the invitation for enrollment in a trial with their partner, where the partner’s opinion influenced the choice on participation. In all but two cases this was a unanimous decision, in the two cases the woman and her partner disagreed. Women indicated that opinions of persons other than their partner were not very influential.

“I discussed it with my husband. I thought, like, if it would have been only my decision, I would have agreed to participate. My answer depended on my husband’s opinion. He thought it was a good decision, so we unanimously agreed on participation. Interviewer: “What if you partner had disagreed? ”Participant: “Then I would not have participated in the trial.”” (Allo trial participant)

Women indicated they had decided on participation without consulting their gynecologist, however when the gynecologist was contacted, his or her opinion was mostly influential. All respondents felt free to make their own decision, without feeling pressure from anyone to participate.

B. (Contribution to) research and healthcare
Women indicated as a reason for participation their contribution to scientific research, as they were convinced about its importance.

“I reasoned also, like, these are studies for the future, and I have a daughter, and you never know… I am prepared to participate for others, so things will be better in the future than how they are now. I am benefiting from what others have done before me.” (Hypitat II trial participant)

Interviewees who had declined participation also judged scientific research important – either mentioned in the open question or during the semi-structured part of the interview. In their case other themes, like uncertainty, the intervention, the trial design and their personal situation, outweighed this importance.
Table 3: Main themes that influence trial participation

<table>
<thead>
<tr>
<th>Theme</th>
<th>Sub codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>External influence</td>
<td>▪ Concern from social environment</td>
</tr>
<tr>
<td></td>
<td>▪ Trust in the health professional</td>
</tr>
<tr>
<td></td>
<td>▪ Feeling of disappointing the health professional</td>
</tr>
<tr>
<td>Research and healthcare</td>
<td>▪ Familiarity with scientific research</td>
</tr>
<tr>
<td></td>
<td>▪ Willingness to contribute to research</td>
</tr>
<tr>
<td></td>
<td>▪ Feeling of participating in an experiment</td>
</tr>
<tr>
<td>Perception one’s own situation</td>
<td>▪ Perception own situation and medical history</td>
</tr>
<tr>
<td></td>
<td>▪ Feeling very eligible or very ineligible for scientific research</td>
</tr>
<tr>
<td>Trial design</td>
<td>▪ Randomization</td>
</tr>
<tr>
<td></td>
<td>▪ Blinding</td>
</tr>
<tr>
<td></td>
<td>▪ Placebo</td>
</tr>
<tr>
<td></td>
<td>▪ Additional efforts</td>
</tr>
<tr>
<td></td>
<td>▪ Insurance of medical research</td>
</tr>
<tr>
<td>Intervention</td>
<td>▪ Intervention</td>
</tr>
<tr>
<td></td>
<td>▪ Natural course</td>
</tr>
<tr>
<td>Information and counseling</td>
<td>▪ Written information</td>
</tr>
<tr>
<td></td>
<td>▪ Counseling: information and timing, atmosphere</td>
</tr>
<tr>
<td></td>
<td>▪ Time for consideration on participation</td>
</tr>
<tr>
<td>Uncertainty</td>
<td>▪ Fear</td>
</tr>
<tr>
<td></td>
<td>▪ Stress</td>
</tr>
<tr>
<td></td>
<td>▪ Doubt</td>
</tr>
<tr>
<td></td>
<td>▪ Physician does not know what is best</td>
</tr>
</tbody>
</table>

A participant of the Ppromexil trial suggested to improve publicity on clinical trials and research in pregnancy:

“Maybe one should increase the awareness about the existence of studies one can participate in in case of pregnancy. Maybe, somehow, more people should be informed once pregnant, so they know about trial participation. Myself, I did not think about it - I have not experienced this before. I think receiving a folder with ‘scientific research for pregnant women’ in advance would decrease the level of stress. If you had read it, you would know it might be coming up. So one can already think about research.”

C. Perception of one’s own situation
The personal perception of one’s own situation appeared influential in the decision on participation: women considered themselves either very eligible or not at all eligible for scientific research, sometimes taking into account their current complicated pregnancy, medical history or their own nature.
“There are people who participate in trials; that is very special and good, but I am not such a person, all that twiddling to my body. Maybe I would if it had been a singleton pregnancy, but now, with twins, it is already scary: and all the twiddling to your body. I prefer nature”. (ProTwin trial non-participant)

Intuitional or emotional aspects seemed to be influential. This became apparent in citations and also when explicitly asked whether their decision was a rational decision, women answered they trusted their feelings or they were inclined to participate but it did not feel good.

D. Trial design
Randomization was perceived as negative by women, which resulted in uncertainty. Women could not explain (in any way) why randomization was used, or could be used, however this lack of knowledge was not necessarily a barrier for participation.

“If I had decided to participate, there would be uncertainty about induction. So the disappointment can be huge. No, with the uncertainty you don’t know if you take a left or a right. If you decide yourself, you know where you go. You now: I go right”. (HYPITAT II trial non-participant)

E. Intervention
Participants mentioned potentially receiving the intervention as a reason for participation, either because of the potential therapeutic benefit, or since they preferred an intervention over expectant monitoring or no intervention.

“Well, if in this case, if they stay in longer, that’s an advantage for me as well. That was actually the only reason to participate. But I needed to be convinced that there were no disadvantages, that it was not detrimental if they stayed in shorter, because of that”. (ProTwin trial participant)

Other women disliked an interventional (“active”) strategy, and rather preferred the natural course, or were more focused on potential (unknown) negative effects. As an explanation for this, women mentioned that the risk of the natural course is one you do not choose for, but which is already present, contrary to the eventual risk of an intervention or trial participation, which feels as the women’s choice and therefore more as their responsibility. All non-participants stated that a negative association with the
intervention or a direct or more indirect negative effect of intervention, as discussed under the main motivations.

“They were uncertain about side-effects for the baby, so then I decided not to take any risk. To me it was already pretty clear: during my time here, I wanted to let mother nature take its course. I am not going to mess with it. If nature decided it to be this way, I let it be, you know.” (ProTwin trial non-participant)

F. Information and counseling

Women considered the information adequate. However, respondents remarked that the counseling was very hastily. A no rush atmosphere, often where counseling was done by a research nurse or midwife, with sufficient time to discuss patients questions was viewed as positive.

“Thinking back, I realized it matters a great deal who comes to inform you about the study. Imagine a research midwife is standing at my bed, taking the time, versus a doctor is sitting at the windowsill, just not looking at her watch, saying “I have 5 minutes, so you have to decide now, otherwise it will be too late.” That makes a difference, and influences the outcome of the decision.” (Apostel II participant)

One women mentioned to have received unclear and incomplete written information, but that did not withhold her from participation.

“I only understood later that these were the same tocolytics as you would receive usually, there is nothing different about. It is not a new medication, that’s what I understood later. That was unclear at the time I had to decide, it seemed if it was a new medication, with a new method to look at whether the baby would stay in longer with premature rupture of membranes and what the harmful effects for the child or the mother would be. If they had explained it better, had told me what the potential adverse effects were - that is of course the point of the trial - than it would have been easier to participate. If they had only said something like “the only potential harmful effect is that you baby may be a bit smaller, or bigger, or more left, or right” but it’s quite difficult if you don’t know. It’s an ethical dilemma.” (Apostel II trial participant)

All but one women judged the time to consider participation adequate, or they understood why the time to consider participation was short (as in the Allo trial). Women
declined participation because there was a too overwhelming amount of new information, or the timing was not very well.

G. Uncertainty
The theme “uncertainty” emerged both in women who accepted and in women who had declined the invitation to participate. Non-participants explicitly mentioned to have declined participation because the feelings of uncertainty, even before they had reached the stage explicitly weighing advantages and disadvantages on participation. Women who declined participation indicated this uncertainty prevailed over other factors that could have led to participation.

“No, I did not consider that, I did not think about it. For me the safety of the baby was most important. No, I did not see an advantage. No, that ‘advantages aspect’, they did not talk about it. And I did not ask for it.” (Promexil trial non-participant)

Both participants and non-participants indicated that the invitation to participate in scientific research was stressful. Being confronted with an (unexpected) choice about trial participation was a decision that needed thorough consideration.

“Whether it is really stress, I am not sure, we have talked about it a lot, both my husband and I, and also with a friend of mine who lives in Rotterdam. It was also on my mind quite a lot, but whether it has caused me physical stress, I am not sure. Yes, I have thought a lot about it, as one never seems to make the right decision. If he had been born, and something had been wrong, while I had not participated, I would have wished I had. On the other hand: if I had participated in the trial and something had gone wrong, I would have wished not to have participated in the trial.” (Apostel II trial participant)

Some of the women were really surprised when confronted with the fact that ‘2010 state of the art health professionals’ do not know what is best.

“They said “We think it is silly to say - and may sound very strange to you - but we have to be honest: we don’t know”. And I was lying there and thinking all the time “I’ll see what happens” until that moment. Then I thought “I feel left to my fate”. I thought it was very honest, but also very hard. You are there for a reason, and they are supposed to know, they have studied for this. I assumed they could tell me in what direction to go, but that, they could not. That is really tough. They could only provide me with certain facts, that neither actually, and the research was there for a reason.” (Hypitat II trial non-participant)
DISCUSSION

Contribution to scientific research was for many participants their main motive for participation in the trial, while others mentioned to have participated because the specific intervention was not available outside the trial. Key motives for non-participation were a negative association or dislike of the intervention, either because it might do harm or for practical reasons. We identified seven themes that influenced trial participation. We noted that uncertainty about scientific research and/or the intervention was reported to be of considerable importance.

This study has looked at a variety of trials, not selected on their recruitment performance, but we selected from running at the time of our study within the Dutch consortium for obstetric studies. We have sampled patients from multiple centers, invited for enrollment by diverse health professionals, in three different geographical areas in the Netherlands. This way we aimed at having a detailed picture of reasons for (non)enrollment in clinical trials.

Fortunately, most of the invited took part in the interview, only two of eleven non-participants invited by phone declined to be interviewed, and one interview had to be cancelled because of medical reasons. We expected that patients who had declined enrollment in the trial would also be more tended to decline an interview, and thereby excluding general reasons for non-participation in health research, unrelated to the specific trial.

A number of potential limitations of our analysis may invite discussion. In general, qualitative research may be seen as vulnerable, since interpretation is an inevitable part of the analysis of the transcripts. This could lead to difference in interpretations between researchers. To reduce this risk, two researchers examined the transcripts. Discrepancies were discussed until agreement was reached.

Five of the 21 respondents were interviewed more than three months after the invitation to enroll in a clinical trial. It is possible that they could recall all factors that influenced the decision, or that their memories differed from their thoughts in the decision making process. All interviewees stated, to our surprise, that the counseling and the decision making process were very well remembered. During the interview only incidental a respondent said not to remember well if a specific topic had been discussed.

The seven themes we identified in this study have been mentioned before in the literature.
Kenyon et al. performed interviews with women who had participated in the ORACLE trial, a randomized trial investigating the value of administration of antibiotics during premature labor.\textsuperscript{18}

They concluded that women gave prominence to the socio-emotional aspects of their interactions with healthcare professionals in making decisions on trial participation. The interviews suggested that the stressful nature of the situation affected their ability to absorb the information. The main motivation for trial participation was the possibility of an improved outcome for the baby. The second motivation was an opportunity to help others, but this was conditional on there being no risks associated with trial participation. McCann and colleagues introduced the term ‘conditional altruism’ based on non-participant observation of recruitment consultations and in-depth interviews with people invited to participate in the UK REFLUX trial. It describes that the willingness to help others that may initially incline people to participate in a trial, but that is unlikely to actually lead to trial participation unless people also recognize that participation will benefit them personally.\textsuperscript{19}

Uncertainty due to unfamiliarity with research or research methods was also identified as a theme related to trial participation in pregnant women by Mohanna et al,\textsuperscript{5} and in a systematic review by Ross et al, not restricted to pregnant women.\textsuperscript{20} Women reported that they would let mother Nature do her work, and were reluctant to actively choose an intervention in what until then was perceived as an uncomplicated pregnancy. Lyerly et al report that risks associated with undertaking medical interventions during pregnancy were focused on, not taking into account the demonstrable risk to both woman and fetus of failing to intervene.\textsuperscript{21}

Unfamiliarity with randomization was a source of uncertainty; for many patients it remained unclear why randomization is used in scientific research. Robinson and colleagues investigated lay public’s understanding of equipoise and randomization in randomized controlled trials from different perspectives.\textsuperscript{22} The research was not carried out in real healthcare settings. Even participants who could correctly explain the rationale behind random allocation methods, judged it as unacceptable. They doubted the possibility of individual equipoise and saw no scientific benefits of random allocation over doctor/patient choice. Robinson et al. concluded that, given the extent of disparity between the assumptions underlying trial design and the assumptions held by the lay public, the solution is unlikely to be simple. Many women were surprised to learn that the
doctor does not know what is best. This was also reported by Mohanna: ‘Some patients will prefer to assume that [My] doctor knows best [about me and my baby], and not be happy to enter into the discussion of uncertainty that a trial and the issue of informed consent will raise’. \(^{23}\) Counseling by research staff, instead of the treating physician, seemed to positively influence participation, which has also been also suggested by the review of Tooher and colleagues. \(^{6}\)

To identify strategies to reduce feelings of uncertainty and stress, further research could elaborate on the work by Junghans and colleagues, where an opt-out versus an opt-in design for low-risk interventions was proposed, as an opt-in system resulted in lower response rates and a biased sample. \(^{24}\) This could not only increase participation rates, but might also shift the responsibility and difficult decision process from pregnant women to the health professionals. Crombie added the following to this discussion: “Research should be undertaken only when there is a high likelihood of producing valid findings. Ethics requirements which result in invalid research may themselves be unethical.” \(^{25}\) Ethical committees should be responsible for determining which trials are eligible to run in this system, with a low or no additional risk to the patient. Patients could be informed about this general policy as soon as they enter the hospital, and are invited to sign a general informed consent about the use of data and efforts to improve quality.

Alternatively, one could think of a classification system of trial risk, where the potential risks of a trial are set out in a uniform label, like energy labels, to make them more transparent for patients. A class A trial, could for example mean that widely used interventions are compared, without additional risk above usual clinical practice. A class E trial could mean that the new intervention is highly experimental. This classification could be proposed by the principal investigator of the trial, and confirmed by an ethical committee before its use. In addition, one could imagine that health professionals recommend participants to participate in Class A trials, of low risk, instead of explicitly leaving the choice to the (vulnerable) patient.

Uncertainty could also be reduced, and awareness improved, when pregnant women become more familiar with scientific research in general, and research in pregnancy in specific. A national public campaign, or an information leaflet introducing the goals, methods and necessity of scientific research when entering a midwifery practice or a
hospital could habituate women to scientific research and the methods used. To do so, for example, in 2008 the ‘Get Randomized’ campaign was launched in Scotland, informing the public about the importance of clinical trials using television, radio and newspaper advertising. It showed an improvement in public awareness of clinical trials following the campaign. However, on whether those who recalled the advertising would personally take part in a clinical trials if invited, there was little difference in response following the campaign.\textsuperscript{26} In the United States a longer running public service advertising campaign celebrating the ‘everyday medical heroes’ of clinical research has been set up because of the believe that the public has a poor and often negative understanding of clinical research.\textsuperscript{27} When patients are aware, or even expecting trials, and regularly ask their health professional whether any trials are running in the department, a clinician might be more inclined to bring up the subject of trial participation, which could have a synergistic effect.

We observed that uncertainty about scientific research and the intervention evaluated was reported to be of considerable importance. A more thorough understanding and knowledge on potential barriers and facilitators may help to improve participation of pregnant women in future trials.
Chapter 5 | Pregnant women’s concerns when invited to a randomized trial

REFERENCES


Website: http://www.ciscrp.org/patient/
Chapter 5 | Pregnant women’s concerns when invited to a randomized trial

APPENDIX 1: TOPIC LIST

Openingsvraag: Wat deed u besluiten om wel/niet mee te doen aan het <trial acronym noemen> onderzoek?

TOPICS

- **Persoonlijk voordeel**: therapeutisch voordeel bij deelname, meest recente therapie, voldoening, meer/betere monitoring van ziekte, betere relatie met behandelaar, vrijheid
- **Onbaatzuchtigheid**: bijdrage wetenschap, behandelaar een plezier doen, andere patiënten met deze aandoening helpen
- **Kennis/informatieverstrekking**: precies weten waar je voor kiest, elk moment kunnen stoppen, randomisatie, placebo, geblindeerd onderzoek/dubbel blind onderzoek, doel van de studie, patiëntinformatie
- **Bezorgdheid/wantrouwen**: nadelige gevolgen behandeling, bekende behandeling beter, (extra) injecties/medicatie, bezorgdheid informed consent, angst voor onbekende, verlies controle, inbreuk privacy, studiedesign, stress
- **Organisatorisch**: afspraken, bedenktijd, extra consulten/injecties, reistijd en kosten, werk, kinderopvang, tijd voor “onderzoeks consult”
- **Attitude**: houding ten opzichte van wetenschappelijk onderzoek
- **Sociale omgeving**: invloed partner/omgeving/internet, wie maakte keuze, behandelaar niet enthousiast over studie, verplichting ten opzichte van behandelaar, moment van counseling
CHAPTER 6

DOES RECRUITMENT FOR MULTICENTER CLINICAL TRIALS IMPROVE DISSEMINATION AND TIMELY IMPLEMENTATION OF THEIR RESULTS?

A SURVEY STUDY FROM THE NETHERLANDS

Rogier J.N.T.M. Litjens, Katrien Oude Rengerink, Nora A. Danhof, Roy F.P.M. Kruitwagen, Ben Willem J. Mol

Clinical Trials, 2013;10:915-23
ABSTRACT

Background Results from clinical trials are often slowly implemented. We studied whether participation in multicenter clinical trials improves reported dissemination, convincement, and subsequent implementation of its results.

Methods We sent a web-based questionnaire to gynecologists, residents, nurses, and midwives in all obstetrics and gynecology departments in the Netherlands. For nine trials in perinatology, reproductive medicine, and gynecologic oncology, we asked the respondents whether they had knowledge of the results, were convinced by the results, and what percentage of their patients were treated according to the results of these trials. We compared the level of knowledge, convincement, and reported implementation of results in practice for the nine trials for respondents who worked in hospitals that had recruited for a trial with respondents who worked in a hospital that had not recruited for that trial. The reported implementation was restricted to six trials that showed decisive results.

Results We analyzed 202 questionnaires from 83 departments in obstetrics and gynecology in the Netherlands (93% of all departments). The percentage of respondents who had worked in a hospital that recruited for a specific study varied between 8% and 71% per study and was 28% on average. The relative risk (RR) for knowledge of the study result for respondents who had worked in a recruiting hospital was for all studies positive and varied between 1.1 and 3.3 (pooled RR: 1.8, 95% confidence interval (CI): 1.7–1.9). In general, health-care workers were convinced of trial results, independent of whether they had worked in a hospital that recruited for a trial or not (pooled RR: 1.02, 95% CI: 0.99–1.05). Reported implementation of trial’s results, that is, less than 20% were treated with unfavorable treatment according to study results, was better in hospitals that had recruited for those trials (pooled RR: 1.1, 95% CI: 1.02–1.19).

Conclusion Participation in these multicenter clinical trials was associated with better knowledge about the trial’s results, with a minor improvement of the reported implementation of the study results.
INTRODUCTION

Health-care decisions should be based on evidence, preferably on high-quality evidence like randomized controlled trials. Generating evidence by performing these trials is expensive and time-consuming. Ironically, one of the most consistent findings in research in health care is the gap between available evidence and practice.\(^1\) Results of studies in the United States and the Netherlands suggest that about one in three patients do not receive care according to the present scientific evidence. And in about one in every four patients, the provided care is not needed or even potentially harmful.\(^2\)\(^3\)

Introduction of new evidence, even evidence from high-level evidence like clinical trials, in routine daily practice remains difficult. Even when research results are successfully disseminated, diffusion of the innovation occurs slowly, if at all.\(^2\) Many obstacles impede implementation of evidence-based interventions in practice.\(^4\)\(^-\)\(^7\) Grol and Grimshaw concluded from an overview of the literature on techniques and strategies to improve implementation that a change in behavior is difficult and generally requires comprehensive approaches at different levels: the level of the patient, the doctor, the team practice, the hospital, and the wider environment.\(^8\) A key to implementation of clinical research into everyday practice is knowing the outcomes of relevant clinical trials. Forsetlund et al. concluded in their 2009 Cochrane review that educational meetings alone, or combined with other interventions, can improve professional practice and health-care outcomes for the patients.\(^4\)

Involvement of health-care workers in the creation of evidence might facilitate dissemination and implementation. This could occur by increasing knowledge of the trial results in departments recruiting for the trial and by increasing the willingness to change according to the results of the trial, as they were willing to randomize patients and have used the intervention themselves. Research conducted in the field of business and marketing in 2011, by Norton et al., shows that labor enhances affection for its results.\(^9\) When people construct products themselves they put greater value on their creations. This phenomenon is referred to as the IKEA effect, named after the Swedish manufacturer whose products typically arrive with some assembly required.\(^10\) We hypothesized that this effect might also play a role in implementation of health-care interventions, wherein caregivers who helped create research knowledge might be more likely to apply it than those who didn’t participate in its creation.
A study by Ketley and Woods in 1993 has shown that the level of use of thrombolytic drugs was strongly associated with the extent of involvement in multicenter trials of thrombolysis: 64% of the variation in thrombolytic use between districts in that year could be accounted for by districts’ participation in multicenter trials during the preceding 2 years (p = 0.003). Shah et al. studied the influence of participation of sites in the Superior Yield of the New Strategy of Enoxaparin, Revascularization and Glycoprotein IIb/IIIa Inhibitors (SYNERGY) trial on the type of heparin subsequently used; site participation did not alter patterns of low-molecular-weight heparin use for non-ST-elevation acute coronary syndromes after publication of the trial results.

Majumdar et al. found that sites that had taken part in the Survival and Ventricular Enlargement (SAVE) trial were no more likely to adopt angiotensin-converting enzyme (ACE) inhibitors for patients with myocardial infarction than were sites that had not taken part. They concluded that if those who generated the evidence are slow to translate it into practice, it is unlikely that passive forms of dissemination can improve the quality of care. While hospitals recruit for studies to generate evidence that might be used globally, local participation might also boost the level of evidence-based practice locally and provide more insight into implementation of trial results in clinical practice.

We constructed a questionnaire to determine whether gynecologists, residents, midwives, and nurses reported better knowledge, were more convinced, and more often implemented results of a clinical trial when they worked in an obstetrics and gynecology department that recruited for that trial.

METHODS

PARTICIPANTS

We asked all 89 departments of obstetrics and gynecology in the Netherlands to participate in our survey, conducted between September 2010 and April 2011. Eight of the 89 departments were academic and 81 were non-academic. We emailed a web-based questionnaire to a mailing list of participants of the Consortium in Obstetrics, Gynecology, and Subfertility in the Netherlands. Due to a low response, we changed the strategy and sent it to two gynecologists in each department and requested to distribute the list to a resident, a midwife, and/or a nurse. To reduce selection bias, we asked the
two gynecologists lowest in alphabetical order on each hospital’s website to fill out the questionnaire. We requested the gynecologists to distribute the questionnaire within their department to two residents and/or nurses if they were not mentioned on the website. As a reminder, two gynecologists, a resident, a nurse, and (if present) a clinical midwife of the departments that had not responded initially were sent a paper version.

As multiple responses from respondents in one department might influence the results, we allowed from each hospital a maximum of two questionnaires per profession group (gynecologists, residents, and nurses or midwives). When more than six questionnaires were returned from one hospital, we included the two that were received last, per profession group, to reduce selection bias.

**QUESTIONNAIRE**

We selected nine randomized clinical trials that were run in a nationwide Consortium for Obstetrics, Gynecology, and Subfertility in the Netherlands and published their results between 2004 and 2010: Hypertension and Pre-eclampsia Intervention Trial at Term (HYPITAT),14 Intra-Uterine Pressure Catheter (IUPC),15 Anticoagulants for Living Fetus (ALIFE),16 External Cephalic Version (ECV),17 Bed Rest,18 Laparoscopic Electrocautery of the Ovaries (LEO),19 Metformin,20 Disproportionate Intra-Uterine Growth Intervention Trial at Term (DIGITAT),21 and the Total Laparoscopic Hysterectomy (TLH)22 (Table 1).

In our questionnaire, we introduced each of the studies and the study results to those surveyed. For each trial, we summarized the conclusion in one sentence. For each of these trials, we asked health professionals (1) whether they had knowledge of the study results previous to information presented in the survey questionnaire; (2) whether they were convinced by the study results (yes/no/unknown); (3) in what percentage of patients they currently performed the clinical procedures tested in the trial in clinical practice (0%-20%, 21%-40%, 41%-60%, 61%-80%, 81%-100%, or unknown); and (4) whether they considered this percentage as just right, too high, or too low.

Respondents were also asked whether or not they had worked in a center that had randomized patients for each of these nine trials. We also collected general characteristics of the respondents including age, sex, (sub) specialization, type of hospital they worked in, and having a PhD.
STATISTICAL ANALYSIS
For the analysis, we compared knowledge of the results (yes/no), whether they were convinced of the study findings (yes/no), and level of reported implementation in practice for the nine trials, between respondents who had worked in hospitals that had or had not recruited for that trial. For assessment of the level of implementation, only studies that indicated a clear recommendation on use or disuse of a treatment were included: HYPITAT, IUPC, ALIFE, ECV, Bed Rest, and Metformin (Table 1). Although the level of implementation was provided by respondents as the percentage of patients treated according to the preferred strategy, we dichotomized implementation as ‘yes’ in case of 0%–20% for unfavorable treatment or 81%–100% for favorable treatments or ‘no’ if 21%–100% implementation for unfavorable treatments and 0%–80% for favorable treatments. If respondents did not know the status of implementation in their clinic, they were excluded from this aspect of the analysis.

Differences in knowledge, convincement, and implementation between hospitals that were recruiting and hospitals that were not recruiting for a trial were shown as relative risks (RRs) with 95% confidence intervals (CIs). To calculate p-values, we used chi-square tests for categorical data, except in cases where the expected cell count was below 5, and then Fisher’s exact test was used.

We planned to perform a multivariable logistic regression analysis in which we incorporated apart from having recruited for a trial also age, sex, having a PhD, and years working in this function as predictors for implementation.

The analysis was performed with SPSS 18.0 for Windows (SPSS, Inc., Chicago, IL). All p-values were two sided and p < 0.05 was considered statistically significant.

RESULTS
RESPONDENTS
We received 286 questionnaires. We excluded 57 questionnaires because they were incomplete due to technical errors in the questionnaire or unknown reasons, as well as 27 questionnaires for exceeding the maximum of two responses per profession group from one hospital. This resulted in 202 questionnaires that were used for analysis. At least one questionnaire was received from 83 of the 89 (93%) obstetrics and gynecology departments in the Netherlands, including all 8 academic hospitals. Five gynecologists
**Table 1:** Results of the included studies and consequences of implementation of studies

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Year</th>
<th>Outcome</th>
<th>Definition of successful implementation</th>
<th>Result of most recent meta-analysis (if available)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEO</td>
<td>2004</td>
<td>Laparoscopic electro cautery equally effective compared to the use of recombinant follicle stimulating hormone in women with the polycystic ovary syndrome</td>
<td>Excluded, no clear preference</td>
<td>Not available</td>
</tr>
<tr>
<td>Metformin</td>
<td>2006</td>
<td>Metformin is not an effective addition to clomifene citrate as the primary method of inducing ovulation in women with polycystic ovary syndrome</td>
<td>In ≤20% use of metformin in addition to clomifene citrate</td>
<td>No evidence to support the use of metformin is an addition to clomifene citrate as the primary method of inducing ovulation in women with polycystic ovary syndrome</td>
</tr>
<tr>
<td>ECV</td>
<td>2008</td>
<td>The use of nifedipine did not increase the rate of successful external cephalic versions compared to placebo</td>
<td>In ≤20% use of nifedipine during external cephalic versions</td>
<td>No evidence to support the use of nifedipine during external cephalic version</td>
</tr>
<tr>
<td>HYPITAT</td>
<td>2009</td>
<td>Induction of labor in women with pre-eclampsia or gestational hypertension at a gestational age of &gt;36 weeks results in a better outcome than expectant management for the woman and a similar outcome for the child</td>
<td>In ≥80% induction of labor in women with pre-eclampsia or gestational hypertension at a gestational age of &gt;36 weeks</td>
<td>Not available</td>
</tr>
<tr>
<td>Bed Rest</td>
<td>2009</td>
<td>Immobilization after intra-uterine insemination results in higher pregnancy rates than immediate mobilization</td>
<td>In ≥80% immobilization immediately after intra-uterine insemination</td>
<td>Not available</td>
</tr>
<tr>
<td>IUPC</td>
<td>2010</td>
<td>The use of an intra-uterine pressure catheter did not significantly improve the outcome during induced or augmented labor</td>
<td>In ≤20% use of intra-uterine pressure catheters during induced or augmented labor</td>
<td>Not available</td>
</tr>
<tr>
<td>ALIFE</td>
<td>2010</td>
<td>The use of anticoagulants in women with unexplained recurrent miscarriage did not increase the ongoing pregnancy rate</td>
<td>In ≤20% use of anticoagulants in women with unexplained recurrent miscarriage</td>
<td>No evidence to support the use of anticoagulants in women with unexplained recurrent miscarriages</td>
</tr>
<tr>
<td>DIGITAT</td>
<td>2010</td>
<td>No clear advantage or disadvantage for induction of labor when disproportionate intra-uterine growth is present</td>
<td>Excluded, no clear preference</td>
<td>Not available</td>
</tr>
<tr>
<td>TLH</td>
<td>2010</td>
<td>No clear advantage for laparoscopic hysterectomy and bilateral salpingo-oophorectomy in women with early stage endometrial cancer</td>
<td>Excluded, no clear preference</td>
<td>Not applicable</td>
</tr>
</tbody>
</table>

HYPITAT: hypertension and pre-eclampsia intervention trial at term; IUPC: intra-uterine pressure catheter; ALIFE: anticoagulants for living fetus; ECV: external cephalic version; LEO: laparoscopic electro cautery of the ovaries; DIGITAT: disproportionate intra-uterine growth intervention trial at term; TLH: total laparoscopic hysterectomy.
and three residents did not report in which hospital they worked. Eight of the 83 hospitals (10%) were academic and 75 (90%) were non-academic. For a 100% response rate, 534 questionnaires would have to have been returned (two gynecologists, two residents, one research nurse, and one midwife for each of the 89 departments). Therefore, the response rate is at least 38% (202/534).

The age of respondents ranged between 23 and 65 years, on average 43 years. A total of 134 (66%) respondents were women; 65 (32%) had a PhD. As a reference, 41% of members of the Dutch Society for Obstetrics and Gynecology have a PhD (525/1285). Overall, 58% were gynecologists, 23% midwives, 17% residents, and 2% research nurses. For the nine trials, the percentage of respondents who had recruited patients for a trial ranged from 8% (Metformin) to 71% (HYPITAT), 28% on average.

Table 2: Knowledge of study outcome versus participation in the study

<table>
<thead>
<tr>
<th>Trial</th>
<th>Recruited for trial</th>
<th>Did not recruit for trial</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Knows outcome</td>
<td>Knows outcome</td>
<td></td>
</tr>
<tr>
<td>HYPITAT</td>
<td>97% (138/143)</td>
<td>86% (51/59)</td>
<td>1.12 (1.00–1.24)</td>
</tr>
<tr>
<td>IUPC</td>
<td>100% (30/30)</td>
<td>77% (133/172)</td>
<td>1.29 (1.19–1.40)</td>
</tr>
<tr>
<td>ALIFE</td>
<td>84% (59/70)</td>
<td>72% (95/132)</td>
<td>1.17 (1.01–1.36)</td>
</tr>
<tr>
<td>ECV</td>
<td>97% (28/29)</td>
<td>43% (75/173)</td>
<td>2.23 (1.85–2.67)</td>
</tr>
<tr>
<td>Bed Rest</td>
<td>100% (23/23)</td>
<td>69% (123/179)</td>
<td>1.46 (1.32–1.61)</td>
</tr>
<tr>
<td>LEO</td>
<td>93% (25/27)</td>
<td>38% (66/175)</td>
<td>2.46 (1.97–3.06)</td>
</tr>
<tr>
<td>Metformin</td>
<td>94% (16/17)</td>
<td>29% (53/185)</td>
<td>3.29 (2.54–4.24)</td>
</tr>
<tr>
<td>DIGITAT</td>
<td>94% (122/130)</td>
<td>61% (44/72)</td>
<td>1.54 (1.27–1.86)</td>
</tr>
<tr>
<td>TLH</td>
<td>92% (36/39)</td>
<td>39% (62/159)</td>
<td>2.36 (1.91–2.93)</td>
</tr>
</tbody>
</table>

HYPITAT: hypertension and pre-eclampsia intervention trial at term; IUPC: intra-uterine pressure catheter; ALIFE: anticoagulants for living fetus; ECV: external cephalic version; LEO: laparoscopic electro cautery of the ovaries; DIGITAT: disproportionate intra-uterine growth intervention trial at term; TLH: total laparoscopic hysterectomy.

Percentages shown are calculated based on the total number of people who did or did not recruit for a specific trial.
KNOWLEDGE OF TRIAL RESULTS
For all nine studies, health professionals in centers which recruited for the trial had significantly more knowledge of the results of the study compared to those that did not recruit for the trial (Table 2). The percentage of respondents with knowledge of the results of the study ranged from 84% to 100% for respondents who worked in a hospital that recruited for a trial versus 38% to 86% for centers that had not recruited for the trial. The RR for knowing the study outcomes when respondents had recruited for a trial as compared to respondents who had not recruited, varied between 1.1 and 3.3 for the nine studies (pooled RR 1.8, 95% CI 1.7–1.9, p < 0.01). Overall, participation in any one of the nine trials resulted in a significantly higher percentage of knowing the study outcomes (pooled RR 1.5, 95% CI 1.3–1.8, p < 0.01).

Gynecologists and residents had more knowledge of trial results than midwives, and nurses had the least knowledge of the results of the trials. These differences in knowledge of the study outcome between gynecologists, residents in training, residents not in training, and midwives or nurses were seen for all studies except for the DIGITALTAT trial. For example, for the HYPITAT trial, 100% of 35 residents (both in training and not in training), 96% of 117 gynecologists, and 89% of 46 midwives had knowledge of the trial results.

BEING CONVINCED OF TRIAL RESULTS
Among respondents who had knowledge of the study results, on average, 94% (range 89%–97%) of the respondents were convinced by the results of the studies (Table 3). There was no difference between health professionals who recruited for the trial and health professionals who did not recruit for the trial regarding being convinced of the study results (pooled RR 1.02, 95% CI 0.99–1.05, p = 0.24). Reasons for not being convinced by the trial results were most often that inclusion criteria were restrictive or that the respondent had knowledge of the outcome despite not having read the actual article carefully.

REPORTED IMPLEMENTATION OF TRIAL RESULTS IN CLINICAL PRACTICE
Reported implementation was higher in centers that had recruited for a specific trial. The percentage of respondents reporting that they adequately implemented the results of the study (treatment of less than or equal to 20% of patients with unfavorable treatments and more than 80% with favorable treatments) ranged from 47% to 90% for respondents who worked in a hospital that recruited for a trial versus 28% to 73% for centers that had not recruited for the trial. The RR varied between 0.78 for the Metformin trial and 1.75 for the HYPITAT trial (pooled RR 1.1 and 95% CI 1.02–1.19, p = 0.01, Table 4). Overall,
participation in one of the nine trials resulted in a slightly higher percentage of reported implementation over all studies (82% versus 74%, pooled RR 1.7 and 95% CI 1.2–2.3, p < 0.01). The respondents were not aware of the current clinical practice in their department in 3%–64% of the cases (Table 4).

**Table 3:** Of people who know the study, percentage convinced by outcome according to whether they recruited for the study.

<table>
<thead>
<tr>
<th>Trial</th>
<th>Recruited for trial</th>
<th>Did not recruit for trial</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Convinced</td>
<td>Convinced</td>
<td></td>
</tr>
<tr>
<td>HYPITAT</td>
<td>92% (119/130)</td>
<td>96% (45/47)</td>
<td>0.96 (0.88–1.04)</td>
</tr>
<tr>
<td>I UPC</td>
<td>100% (30/30)</td>
<td>93% (114/122)</td>
<td>1.07 (1.02–1.12)</td>
</tr>
<tr>
<td>ALIFE</td>
<td>96% (54/56)</td>
<td>98% (85/87)</td>
<td>0.99 (0.93–1.05)</td>
</tr>
<tr>
<td>ECV</td>
<td>100% (28/28)</td>
<td>96% (65/68)</td>
<td>1.05 (0.99–1.10)</td>
</tr>
<tr>
<td>Bed Rest</td>
<td>100% (20/20)</td>
<td>91% (102/112)</td>
<td>1.10 (1.04–1.16)</td>
</tr>
<tr>
<td>LEO</td>
<td>96% (24/25)</td>
<td>92% (54/59)</td>
<td>1.05 (0.94–1.17)</td>
</tr>
<tr>
<td>Metformin</td>
<td>93% (14/15)</td>
<td>89% (39/44)</td>
<td>1.05 (0.89–1.25)</td>
</tr>
<tr>
<td>DIGITAT</td>
<td>97% (113/117)</td>
<td>94% (32/34)</td>
<td>1.03 (0.94–1.12)</td>
</tr>
<tr>
<td>TLH</td>
<td>94% (30/32)</td>
<td>93% (55/59)</td>
<td>1.01 (0.9–1.13)</td>
</tr>
</tbody>
</table>

HYPITAT: hypertension and pre-eclampsia intervention trial at term; I UPC: intra-uterine pressure catheter; ALIFE: anticoagulants for living fetus; ECV: external cephalic version; LEO: laparoscopic electro cautery of the ovaries; DIGITAT: disproportionate intra-uterine growth intervention trial at term; TLH: total laparoscopic hysterectomy.

Percentages shown are calculated based on the number of people who did or did not recruit for a specific trial.

Note that the number of respondents does not add up to the amount of respondents in table 2 since some respondents stated to have knowledge of the study results in the first question but stated to be unaware of (the details of) the results in the second question.

When looking at reported implementation rates of specific studies, this difference was statistically significant for the HYPITAT trial (RR 1.8, 95% CI 1.2–2.4), the I UPC trial (RR 1.5, 95% CI 1.3–1.8, p < 0.01), and the Bed Rest trial (RR 1.5, 95% CI 1.1–1.5, p < 0.05). Table 5 shows the relation between knowledge of the study results and reported implementation of its results in practice. The RR for reported implementation of trial results in practice was in favor of the respondents who had knowledge of the trial’s results, with the exception of the ECV and the Metformin trial. It ranged from 0.8 to 2.5 (pooled RR: 1.0, 95% CI: 0.9–1.1,
p = 0.90). For both the ALIFE study and the Bed Rest study, knowledge of study results significantly improved reported implementation of its results in practice. For the HYPITAT study and the IUPC study, the RR was positive but not statistically significant.

**Table 4:** Consistency of therapy with study results by people who recruited or did not recruit for the study

<table>
<thead>
<tr>
<th>Trial</th>
<th>Recruited for trial</th>
<th>Did not recruit for trial</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% of responders</td>
<td>Unaware of</td>
</tr>
<tr>
<td></td>
<td>treating consistent</td>
<td>implementation</td>
</tr>
<tr>
<td></td>
<td>with study results^a</td>
<td>rates</td>
</tr>
<tr>
<td></td>
<td>Unaware of</td>
<td>% of responders</td>
</tr>
<tr>
<td></td>
<td>treating consistent</td>
<td>implementing</td>
</tr>
<tr>
<td></td>
<td>with study results^b</td>
<td>rates</td>
</tr>
<tr>
<td>HYPITAT</td>
<td>73% (95/131)</td>
<td>8% (12/143)</td>
</tr>
<tr>
<td>IUPC</td>
<td>93% (27/29)</td>
<td>3% (1/30)</td>
</tr>
<tr>
<td>ALIFE</td>
<td>92% (49/53)</td>
<td>24% (17/70)</td>
</tr>
<tr>
<td>ECV</td>
<td>95% (20/21)</td>
<td>28% (8/29)</td>
</tr>
<tr>
<td>Bed Rest</td>
<td>94% (17/18)</td>
<td>22% (5/23)</td>
</tr>
<tr>
<td>Metformin</td>
<td>62% (8/13)</td>
<td>24% (4/17)</td>
</tr>
<tr>
<td></td>
<td>RR (95% CI)</td>
<td></td>
</tr>
<tr>
<td>HYPITAT</td>
<td>1.75 (1.24–2.44)</td>
<td></td>
</tr>
<tr>
<td>IUPC</td>
<td>1.50 (1.28–1.76)</td>
<td></td>
</tr>
<tr>
<td>ALIFE</td>
<td>1.10 (0.98–1.23)</td>
<td></td>
</tr>
<tr>
<td>ECV</td>
<td>1.08 (0.96–1.21)</td>
<td></td>
</tr>
<tr>
<td>Bed Rest</td>
<td>1.24 (1.07–1.45)</td>
<td></td>
</tr>
<tr>
<td>Metformin</td>
<td>0.78 (0.50–1.22)</td>
<td></td>
</tr>
</tbody>
</table>

HYPITAT: hypertension and pre-eclampsia intervention trial at term; IUPC: intra-uterine pressure catheter; ALIFE: anticoagulants for living fetus; ECV: external cephalic version; RR: relative risk; CI: confidence interval.

Percentages shown are calculated based on the total number of people who did or did not recruit for a specific trial, ‘% consistent with study results’ excluded responders unaware of implementation rates.

^a** Only studies that showed a clear preference for either using or not using the intervention were included.

^b** Practice by a responder was considered to be consistent with study results if <=20% of patients were treated with unfavorable treatments and >80% with favorable treatments.

Reported implementation of the results did not differ significantly between gynecologists, residents, and midwives for the ALIFE study, ECV study, Bed Rest study, and IUPC study. For the HYPITAT study, residents most often reported implementation of the results in practice: 79% of 29 residents, 66% of 111 gynecologists, and 45% of 40 midwives reported implementation of the results in practice.

To determine whether respondents planned on further implementing the results of the trial, we asked whether they thought the current percentage of implementation was just right, too low, or too high. Of the 215 respondents who reported not implementing the results of the trials, 35% felt that more implementation of the trial’s results was needed. Of the respondents who reported implementing the results of the trials, 5% felt that this level was too low and further implementation was needed. There was no statistically significant difference in the value of the implementation rate between those who had recruited and those who had not recruited for the six trials combined (p = 0.30), nor when looking at each trial individually (p-values not shown).
Table 5: Consistency of therapy with study results by respondents who knew or did not know the study results

<table>
<thead>
<tr>
<th>Trial</th>
<th>Knew the study outcome</th>
<th>Did not know the study outcome</th>
<th>RR (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HYPITAT</td>
<td>65% (115/178)</td>
<td>33% (2/6)</td>
<td>1.94 (0.62–6.04)</td>
</tr>
<tr>
<td>IUPC</td>
<td>67% (104/155)</td>
<td>66% (19/29)</td>
<td>1.02 (0.77–1.36)</td>
</tr>
<tr>
<td>ALIFE</td>
<td>91% (125/137)</td>
<td>53% (8/15)</td>
<td>1.71 (1.06–2.75)</td>
</tr>
<tr>
<td>ECV</td>
<td>88% (80/91)</td>
<td>91% (63/69)</td>
<td>0.96 (0.87–1.07)</td>
</tr>
<tr>
<td>Bed Rest</td>
<td>82% (94/114)</td>
<td>33% (4/12)</td>
<td>2.47 (1.11–5.53)</td>
</tr>
<tr>
<td>Metformin</td>
<td>71% (37/52)</td>
<td>85% (23/27)</td>
<td>0.84 (0.67–1.05)</td>
</tr>
</tbody>
</table>

HYPITAT: hypertension and pre-eclampsia intervention trial at term; IUPC: intra-uterine pressure catheter; ALIFE: anticoagulants for living fetus; ECV: external cephalic version; RR: relative risk; CI: confidence interval.
Percentages shown are calculated based on the total number of people who had or did not have subjective knowledge of the trial results.

*aPractice by a responder was considered to be consistent with study results if <=20% of patients were treated with unfavorable treatments and >80% with favorable treatments.

DISCUSSION

Respondents who had worked in a hospital that recruited for a trial had more knowledge of the trial result, and they more often implemented its results in clinical practice. However, although the RR of the implementation was statistically significant in favor of those working in a hospital that had recruited for a trial, the RR was only 1.1 (95% CI: 1.02–1.19). Convincingness of trial results was high and comparable in both groups. Respondents who had knowledge of the study more often implemented its results in clinical practice, but knowing the results of a study did not result in the same improvements in reported implementation as recruiting for a trial. Those who recruited for the trial know the trial – at least its inclusion criteria – very well and have more insight into the research process.

These findings might be partly explained by the ‘I designed it myself’ or IKEA effect: when persons are involved in creating the evidence, they might be more willing to adopt the evidence. Hypothetically, the quality of care in these hospitals might be improved by participation in (some) trials, as the hospital staff becomes more familiar with and is more aware of the cycle of research and quality improvement. Our study showed a limited effect of recruiting on implementation of evidence in practice.
Two studies in the field of cardiovascular research were unable to find an effect of recruitment for a clinical trial on the implementation rate of the trial results.\textsuperscript{12,13} This might be explained by the fact that the clinicians were not convinced by the results of a single trial. The latter is supported by a study by Ketley and Woods,\textsuperscript{11} who did find that participation in a clinical trial positively influenced the implementation rates. This increase in the implementation rate followed publication of several other studies with similar results to the one clinicians participated in. Knowledge of the trial’s outcome was not measured in any of the trials mentioned. However, other interventions such as participating in a journal club have also been shown to increase the evidence-based knowledge while having only a small effect on actual evidence-based behavior in clinical practice.\textsuperscript{26,27} This might be because other barriers, such as time constraints and the lack of skills for evaluating research results, might restrain health professionals from implementing the results into practice.\textsuperscript{28,29}

A Cochrane review concluded from eight studies evaluating the effectiveness of different interventions designed to support the uptake of systematic review evidence that mass mailing a printed bulletin which summarizes systematic review evidence may improve evidence-based practice when there is a single clear message, if the change is relatively simple to accomplish, and there is a growing awareness by users of the evidence that a change in practice is required. If the intention is to develop awareness and knowledge of systematic review evidence and the skills for implementing this evidence, a multifaceted intervention that addresses each of these aims may be required, though there is insufficient evidence to support this approach.\textsuperscript{30}

A strength of this study is that at least one respondent of 93\% of all hospitals in the Netherlands completed the questionnaire. To ensure representation of all hospitals equally, not more than six questionnaires per hospital were used in the analysis. This number might be viewed as low, as the different people working in the hospitals might have different opinions about implementation of trial results, but most hospitals have their own protocols, based on which we assume responses will (or should) be quite similar within a clinic. We systematically excluded some respondents so as not to over represent certain hospitals, recognizing that it is not optimal to exclude already collected data. The consistency between the results of the studies indicates that – although all run within the Dutch Consortium for Obstetrics, Gynecology, and Reproductive Medicine – the results probably also apply to studies performed in other fields of medicine. These nine studies
were each coordinated by different hospitals and conducted in a different number and different types of hospitals.

There are some limitations of this study. First, we did not have data on implementation rates of the different treatments before start of the trials. We assume the difference in effect between the hospitals will be caused by recruiting for that study, but centers which recruited for the trial might have already been more willing to change their practice and therefore consented to recruit for the trial. Although we have seen an association, we cannot conclude from this cross-sectional study that there is a causal relation between recruiting and knowledge and implementation. It might be that persons who are more research-oriented move to centers that recruit for studies and practice evidence-based medicine. Indeed, a higher percentage of academics recruited patients to the trials compared to the non-academics (45% vs 27%).

Second, we defined the desirability of implementation of trial result rates based on the outcome of these trials only, while implementation should more appropriately be based on high-quality systematic reviews and meta-analyses for practice to be evidence based. Table 1 shows results from the individual trials in our study and meta-analyses of existing trials for similar questions. When meta-analyses were available, they supported implementation of the trial results. At the time when the questionnaire was sent, only the LEO and Metformin trial results had been incorporated in the Dutch guideline but not the findings from the other trials.

A third limitation of our study is that both knowledge of the study results and implementation of the study results in practice are self-reported. Mentioning the results in the questionnaire might have led to socially desirable answers, overestimating the knowledge of studies and the implementation rate, probably especially in hospitals that recruited for the study. A study by Van der Tuuk et al. used the Perinatal Registry in the Netherlands to study the increase in the number of inductions after the HYPITAT trial. They observed a 12.1% increase in the number of inductions in hospitals that recruited, versus a 5.1% increase in the hospitals that did not recruit (p < 0.001). This supports the finding that the implementation rate is higher in hospitals that had recruited for the trial. However, they also observed that the percentage of labor inductions was already higher before start of the HYPITAT trial in hospitals that did not participate. Van der Tuuk et al. showed that in the hospitals that had recruited for a trial, in 66% of eligible women, labor was induced; in
our study, 66% of the respondents stated that in their clinic, in at least 80% of the patients, labor was induced. This suggests that socially desirable answers seem a minor issue.

Fourth, we defined implementation as performing an effective intervention in at least 80% of the cases and performing an ineffective or harmful intervention in a maximum of 20% of the cases. As not practicing according to trial results might have different consequences regarding clinical outcomes and costs, we used this definition to leave some room for exceptions. However, this definition of implementation leaves no room for partial implementation, that is, more often treating patients according to the trial’s outcome, albeit not in 80% of the cases.

Fifth, acquiring knowledge and changing behavior takes time. The included trials ended at different time points, between 2004 and 2010. The available time for implementation and acquiring knowledge about the trials differed from 6 months to 7 years. Nevertheless, participation in a trial resulted in significantly more knowledge of the outcome of all nine trials. Moreover, when asked whether the current implementation rate was ideal, too high, or too low, we did not see a difference between centers that had or had not recruited for the trial.

CONCLUSION
Actively involving health-care workers in the creation of evidence by recruiting for collaborative health research increases knowledge of the trial’s results. However, only a minor effect on the implementation of its results in clinical practice was found, possibly due to the high implementation rates of the six trials.

ACKNOWLEDGEMENTS
We would like to thank all the gynecologists, residents, midwives, and research nurses who responded, for their participation in this study.
REFERENCES


PART II

IMPROVING EVIDENCE BASED CLINICAL PRACTICE
WHO TEACHES THE EVIDENCE-BASED MEDICINE TEACHER?

LETTER TO THE EDITOR

K. Oude Rengerink, K.S. Khan, A.R. Horvath, B. Meyerrose, J. Walczak, K. Suter, B.W. Mol, for EU EBM Unity

Medico Teacher, 2012;34:866
Dear Sir,

There has been recognition of the need to train clinicians in the methods of Evidence-based Medicine (EBM). There is considerable variation in the methods of teaching EBM in clinical settings, very often left to the initiative of enthusiastic individuals. Teach the teacher (TTT) EBM courses might help to improve teaching quality and set a standard for EBM teaching.

We determined the availability and content of TTT EBM courses. A questionnaire developed by the EU EBM partnership was sent to medical institutions potentially offering TTT EBM courses in the United Kingdom, Germany, Hungary, Poland and the Netherlands. The country specific infrastructure and existing networks were used to reach out to target organizations between August 2008 and March 2009. We encouraged recipients to forward the questionnaire if not addressed correctly. All courses of which the organisers stated to be a TTT EBM course were included.

We identified 16 courses out of 114 responses (multiple responses per institution): four in Hungary, four in the Netherlands, four in the United Kingdom and four in Germany. The courses mainly targeted academic specialists or medical practitioners and were taught by academic specialists (N=15) and/or methodologists/statisticians (N=11). Twelve courses regarded teaching the various steps of EBM; seven regarded teaching integration of EBM into daily clinical practice. The clinical learning opportunities chosen were (in order of frequency): journal club, ward rounds, outpatient practice, formal clinical meeting and formal assessments. In eight courses, e-learning modules aided teaching. Eleven courses were officially certified or CME-credited, also eleven concluded with a formal assessment.

Course organisers highlighted the need for help providing practical examples on successful techniques for teaching EBM in clinical practice (N=7), help with a curriculum for trainers to train healthcare workers (N=7) and help with funding of TTT courses (N=7).

The low availability of TTT courses in Europe might indicate a need for development of TTT courses that can be used as a reference point for EBM teaching for postgraduate teachers.

REFERENCE

CHAPTER 8

HOW TO CONFIDENTLY TEACH EBM ON FOOT

DEVELOPMENT AND EVALUATION OF A WEB-BASED E-LEARNING COURSE

Tobias Weberschock, Olanrewaju Sorinola, Shakila Thangaratinam, Katrien Oude Rengerink, Theodoris N. Arvanitis, Khalid S. Khan, EBM Unity Group*

Evidence Based Medicine 2013;18:170-2

*Further members of the EBM Unity Project are: Susanne Weinbrenner, Andrea R Horvath, Rita Onody, Gianni Zanrei, Regina Kunz, Katja Suter, Jacek Walczac, Bernard Burnand, Chantal Arditi, Javier Zamora, Ben WJ Mol, Gemma Barnfield, Harry Gee, Anna Kaleta
ABSTRACT

Background Scarcity of well-trained clinical tutors is a key constraint in integrating teaching of evidence-based medicine (EBM) into clinical activities.

Objectives We developed a web-based educational course for clinical trainers to confidently teach EBM principles in everyday practice. Its e-learning modules defined the learning objectives and incorporated video clips of practical and effective EBM teaching methods for exploiting educational opportunities in six different clinical settings.

Methods We evaluated the course with clinical tutors in different specialties across six European countries using a questionnaire to capture learning achievement against preset objectives.

Results Among 56 tutors, 47 participants (84%) improved their scores from baseline. The mean pre-course score was 69.2 (SD=10.4), which increased to 77.3 (SD=11.7) post-course (p<0.0001). The effect size was moderate with a Cohen’s D of 0.73.

Conclusions An e-learning approach incorporating videos of applied EBM teaching and learning based on real clinical scenarios in the workplace can be useful in facilitating EBM teaching on foot. It can be integrated in the continuing professional development programs for clinical trainers.
INTRODUCTION

Evidence-based Medicine (EBM), an approach to clinical practice, encourages looking up and appraising research evidence for solving patient’s problem instead of just following the wisdom handed down by seniors. Over time, EBM has evolved into a powerful tool for well-informed decision-making integrating the best research evidence with clinical expertise and patient values. Since EBM’s inception, its principles have been taught in undergraduate, postgraduate and continuing education, but incorporating teaching into everyday clinical practice has always been a challenge.

Clinically integrated EBM teaching improves knowledge, skills, attitude and behaviour, while the effect of traditional standalone teaching does not extend beyond changes in knowledge. Despite the recognition that clinical teachers play a key role in the dissemination of EBM, training in how to confidently teach EBM in the workplace has not been widespread in continuing professional development. In New Zealand for example, 40% of the specialists and general practitioners teach EBM, but of these only 68% had formal training in the delivery of EBM teaching. We developed and evaluated a web-based course for training clinical teachers to deliver teaching of EBM alongside service delivery.

COURSE DEVELOPMENT AND EVALUATION

DESCRIPTION OF THE EBM TEACHING-THE-TEACHERS COURSE (http://ebm-unity.pc.unicatt.it/)
Funded by the European Union Leonardo da Vinci Vocational Training Action Programme (project grant number UK/05/B/F/PP-162_349), we developed a web-based course on how to impart clinically relevant EBM teaching in various clinical environments, with input from experienced EBM teachers, clinical epidemiologists, clinicians and educationalists from institutions in seven European countries. An independent steering committee provided input into the process.

The course targets clinical trainers who already possess basic EBM knowledge and skills. The e-learning format allows clinical teachers to undertake the course in the workplace during short breaks between clinical activities. The curriculum defines specific aims and objectives for learning how to teach EBM by exploiting educational opportunities...
in six different clinical settings. The course consists of six e-learning modules that cover EBM teaching on the acquisition, appraisal and application of findings from research in various clinical settings: ward rounds, journal club, audit, outpatient clinics, formal clinical meeting, formal clinical assessment of trainees and audit (Table 1). Audit in this learning context refers to a formal hospital meeting where current clinical practice is evaluated against predetermined standards of care.

In all modules, the key learning objectives are for the teachers to learn how to expose knowledge gaps in their trainees, to lead them to construct structured questions, to help them to track and appraise evidence, and to demonstrate how clinical judgment, patient preferences and research evidence are amalgamated for patient care. Each module lasts about 15 minutes, with the flexibility to stop and restart at the convenience of the participant. The e-learning modules have video clips that demonstrate practical and effective methods of EBM teaching in everyday clinical practice.

COURSE EVALUATION AND OUTCOME MEASURES (ASSESSMENTS)
We evaluated the course in Germany, Hungary, the Netherlands, Poland, Switzerland and the UK. The course was delivered to clinical teachers, from different specialties, in the participating countries. The participants were those clinicians who regularly taught EBM to their postgraduate trainees, within the context of practice. Teachers who only taught EBM in lectures or standalone courses were excluded, as were teachers who lacked basic EBM knowledge. Informed consent was obtained from the participants prior to the administration of the e-learning modules and the completion of the questionnaire instrument. The responses were anonymously analysed.

A before-and-after study design was used to examine the effect of the course on knowledge. The study was completed within a two-month frame. For outcome assessment, we developed and validated a questionnaire to assess the participants’ performance before and after completion of the course. The questionnaire, consisting of 26 multiple choice questions, was used to assess the knowledge of the participants in EBM teaching. The questions posed, reflected the contents of the course’s six e-modules. Their validation involved measurement of Cronbach’s α, inter-item correlation and item discrimination. The questionnaire can be obtained from the authors (RKunz@uhbs.ch).
Participants completed the questionnaire before and after the e-learning sessions. The maximum possible score was 100 points. Pre- and post-course scores were compared, using the paired t-test. We computed the change in knowledge; a positive difference meant a gain in knowledge. The percentage of participants with gain in knowledge and the

Table 1: An outline of the web-based educational course for clinical trainers to confidently teach evidence based medicine (EBM) on foot in everyday practice.

<table>
<thead>
<tr>
<th>Module</th>
<th>Ask</th>
<th>Acquire</th>
<th>Appraise</th>
<th>Assess</th>
<th>Apply</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Ward Rounds</td>
<td>+++</td>
<td>+</td>
<td>+++</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>2. Journal Club</td>
<td>+</td>
<td>++</td>
<td>+++</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>3. Clinical assessment</td>
<td>++</td>
<td>+</td>
<td>++</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>4. Outpatients Clinic</td>
<td>+++</td>
<td>+</td>
<td>++</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>5. Formal Clinical Meeting</td>
<td>+++</td>
<td>+</td>
<td>+</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>6. Audit</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>+++</td>
</tr>
</tbody>
</table>

Assessment: Knowledge questionnaire and reflection using a log of EBM teaching activities

Adapted from Thangaratinam et al.\textsuperscript{11}

Cohen’s d were also computed. Cohen’s d is a measure of the effect size computed as the mean difference of the two interventions over the common standard deviation in order to describe the relative effect in relation to its precision. Cohen refers to a value of 0.2 to small, 0.5 to medium and 0.8 to large effect sizes.\textsuperscript{12}
RESULTS

There were 80 participants, from different specialties, who enrolled in the course. Of these, 70 participants completed the pre-course questionnaire and 56 completed both the pre- and post-course questionnaire. There were two main reasons given for non-completion; time pressure and reported difficulties on accessing the web-based material, in some countries. Participants who completed both questionnaires came from Hungary (14), the United Kingdom (12), Switzerland (11), the Netherlands (9), Germany (7) and Poland (3). Among these 56 teachers, the median age was 45 years, 39% were female and 89% were affiliated to a teaching hospital. We found the mean score for the pre-course assessment to be 69.2 (SD=10.4), while the mean score post-course to be 77.3 (SD=11.7). This improvement (8.1 points) corresponded to a 26% of the maximal possible improvement. Figure 1 shows the distribution of the pre- and post-course scores achieved by the participants. Overall, 47 participants (84%) improved their score between the precourse and postcourse assessment. The paired comparison, between preassessment and postassessment, renders a statistically significant difference (8.1 points; t-test p<0.0001). The magnitude of the effect size, Cohen’s d was 0.73. A trend towards knowledge gain was apparent across different countries and settings, although due to small sample size subgroup analyses did not show statistical significance.

Figure 1: Pre- and post-course scores in a knowledge test among clinical trainers taking the web-based educational course for learning how to teach EBM principles in everyday practice.
COMMENTS

We have developed an on-line EBM teaching-the-teachers course which improves teachers’ knowledge relating to the teaching of EBM in various clinical settings, including ward rounds, audit meetings, journal clubs and outpatient clinics. We have also demonstrated that web-based courses for teaching EBM teachers can be harmonised for delivery across different countries.

One limitation of our evaluation is the relatively small number of participants. The before-and-after design allowed us to initially efficiently pilot the effect of educational materials. It reassures that web-based learning is a feasible method for continuing education in this area. We have previously shown that e-EBM is feasible in postgraduate education in different languages, educational settings, medical disciplines and countries. The absence of a control group in our current evaluation can be seen as a limitation of this study. The assessments before the course served as a baseline control for each individual. Cook et al. have shown that effects are consistent across designs in research on internet-based interventions. Because the before-and-after evaluations were conducted over a short period, the effect of external influences is likely to be negligible and knowledge gains can reasonably be attributed to the effect of exposure to the e-learning course. All participants were already teaching EBM, so their baseline score was high, which makes an increase in the score more difficult to achieve. However, the mean increase was moderate and statistically significant. Although 9 participants showed a drop in scores post-course, 47 (84%) showed knowledge gain. The Cohen’s d of 0.73 suggests a moderate effect size, when compared to other educational interventions.

Evaluation in other settings and with different study designs, particularly with larger numbers, may be undertaken in the future for which the course is made available freely on the Evidence-based Medicine resources page (http://ebm.bmj.com/site/resources/index.xhtml or from http://ebm-unity.pc.unicatt.it/).

In conclusion, a web-based educational intervention on how to teach EBM within the clinical context can be implemented across different countries and in various settings. Empowering EBM clinical teachers this way can pave the way for incorporating EBM into practice through teaching on foot during clinical working hours, rather than standalone teaching off-site.
ACKNOWLEDGEMENTS
Further members of the EBM Unity Project are: Susanne Weinbrenner, Andrea R Horvath, Rita Onody, Gianni Zanrei, Regina Kunz, Katja Suter, Jacek Walczak, Bernard Burnand, Chantal Arditi, Javier Zamora, Ben WJ Mol, Gemma Barnfield, Harry Gee, Anna Kaleta. Contact addresses for the further members of the EBM Unity Project can be found as an additional data file under http://ebm.bmj.com. Address for obtaining the questionnaire from authors: RKunz@uhbs.ch

FUNDING
This study is part of the European Union Evidence-based Medicine Unity (EU EBM Unity) project funded by the European Union Leonardo da Vinci Vocational Training Action Programme (project grant number UK/05/B/F/PP-162_349).
REFERENCES

CHAPTER 9

HOW CAN WE TEACH EBM IN CLINICAL PRACTICE?

AN ANALYSIS OF BARRIERS TO IMPLEMENTATION OF ON-THE-JOB EBM TEACHING AND LEARNING

Katrien Oude Rengerink, Shakila Thangaratinam, Gemma Barnfield, Katja Suter, Andrea R. Horvath, Jacek Walczak, Anna Wetmińska, Susanne Weinbrenner, Berit Meyerrose, Theodoros N. Arvanitis, Rita Onody, Gianni Zanrei, Regina Kunz, Chantal Arditi, Bernard Burnand, Harry Gee, Khalid S. Khan, Ben W.J. Mol

Medical Teacher, 2011;33:e125-30
ABSTRACT

Introduction Evidence-based medicine (EBM) improves the quality of health care. Courses on how to teach EBM in practice are available, but knowledge does not automatically imply its application in teaching. We aimed to identify and compare barriers and facilitators for teaching EBM in clinical practice in various European countries.

Methods A questionnaire was constructed listing potential barriers and facilitators for EBM teaching in clinical practice. Answers were reported on a 7-point Likert scale ranging from not at all being a barrier to being an insurmountable barrier.

Results The questionnaire was completed by 120 clinical EBM teachers from 11 countries. Lack of time was the strongest barrier for teaching EBM in practice (median 5). Moderate barriers were lack of requirements for EBM skills and a pyramid hierarchy in health care management structure (median 4). In Germany, Hungary and Poland reading and understanding articles in English was a higher barrier than in the other countries.

Conclusion Incorporation of teaching EBM in practice faces several barriers to implementation. Teaching EBM in clinical settings is most successful where EBM principles are culturally embedded and form part and parcel of everyday clinical decisions and medical practice.
INTRODUCTION

The amount of medical knowledge is growing exponentially, but integrating research into practice is slow.\textsuperscript{1,2} Health professionals often fail to implement clinical manoeuvres that have established efficacy.\textsuperscript{3} As a consequence patients might receive suboptimal treatment. To stay up to date and deliver optimal health care, health care professionals need to incorporate life-long learning in their profession.\textsuperscript{4} Evidence Based Medicine (EBM) equips doctors with skills to integrate evidence from research in clinical decision making and improves the quality of health care. Professional organisations therefore increasingly promote training in EBM for all health care professions at all levels of education.\textsuperscript{5-7} It has been shown that clinically integrated teaching of EBM is the best way to improve evidence based behavior in practice.\textsuperscript{8} Unfortunately, integration of EBM teaching for postgraduate junior doctors in everyday clinical practice is uncommon and remains a challenge.\textsuperscript{9,10} Courses on how to teach EBM in practice are (scarcely) available in Europe.\textsuperscript{11} Improving knowledge about how best to teach EBM does not automatically lead to implementation of good teaching and learning practice.\textsuperscript{8}

Many previous studies focused on attitudes and barriers for implementing EBM in health care practice,\textsuperscript{12,13} but barriers for implementing the teaching of EBM in clinical practice have only been studied briefly. These barriers and facilitators are currently not well understood. They might differ within and between countries, as they might be related to health care organizational culture, language and availability of evidence and resources to find the evidence.\textsuperscript{14-18}

In this article, we aim to identify and compare barriers and facilitators for teaching EBM in clinical practice in European countries of varying backgrounds. This may provide opportunities for improved strategies of teaching and practicing EBM, which, ultimately, may lead to higher quality and effectiveness of healthcare delivered to patients. It also provides the opportunity to diminish differences in EBM teaching between countries by tackling joint barriers collectively.
METHODS

We conducted a questionnaire survey. Based on literature review in PubMed and input from experts in EBM teaching participating in the EU EBM TTT project (www.ebm-unity.org) a questionnaire was constructed and tested listing potential barriers and facilitators for EBM teaching in clinical practice. The questionnaire also collects demographic characteristics as well as information about how often participants taught EBM in a clinical setting in the last month (questionnaire available upon request). The survey targeted senior clinicians who teach EBM on-the-job in a clinical setting and explored whether they perceive a certain issue as a barrier or facilitator for their teaching. They provided answers ranging from not at all to an insurmountable barrier or facilitator, on a 7 point Likert scale. Issues included attitude, available time, hospital hierarchy, level of understanding English literature, availability of resources, EBM knowledge and skills of teachers, requirements for EBM teaching in curricula or at workplace and availability of Teaching the Teacher courses.

The questionnaire was distributed to participants taking part in an e-learning course Teaching the Teacher, a EU-EBM project funded by the EU Leonardo da Vinci program (www.ebm-unity.org). It was also distributed to EBM teachers participating in a validation study of an assessment tool for this course, and to members of the steering committee of our project. We additionally distributed the questionnaire at an international conference for teachers and developers of EBM (Oxford, December 8-9, 2008). Median scores on the 1-7 Likert scale were used to report the level of being a barrier with 1=not a barrier, 2=very mild barrier, 3=mild barrier, 4=moderate barrier, 5=severe barrier, 6=essential barrier, 7=insurmountable barrier. For facilitators responses were scored as 1=not at all relevant, 2=may be important, 3=slightly important, 4=moderately important, 5=important, 6=very important, 7=essential.

Barriers and facilitators were analyzed over all participants and additionally explorative analyses were stratified per country. Differences between countries were tested using the non-parametric Kruskal-Wallis test or the Wilcoxon Rank Sum test. Differences within countries were tested using the Related Samples Wilcoxon Signed Rank Test. P-values less than 0.05 were considered statistically significant. SPSS version 16.02 was used for all analyses.
RESULTS

A total of 120 clinical EBM teachers from 11 predominantly European countries completed an online or paper questionnaire: 29 from the United Kingdom, 21 from Hungary, 18 from Switzerland, 18 from the Netherlands, 17 from Germany, 4 from Poland and 11 from a variety of other countries (Italy, Belgium, Canada, Finland, Greece and USA). All the 74 participants of the pilot project filled out the questionnaire, all 24 participants of validation of the pilot project, 17 conference participants (small percentage) and 5 members (50%) of the steering committee filled out the questionnaire. Participants of the pilot project filled in an electronic version, all others answered on paper.

Of the participants, 82 were male and 38 female. Thirty-seven were under 40 years old, 78 between 40 and 59 and 4 older than 60 years. Almost all teachers worked in a teaching hospital (N=103, 91%). Fifty-one (43%) teachers stated to teach EBM to postgraduates, 20 (17%) to teach EBM to undergraduates, 20 (23%) to both post- and undergraduates and 20 (17%) stated not to teach EBM in clinical practice.

The most frequently used EBM teaching activities in the last month were demonstration of an electronic search of literature or search strategy and attending a journal club or an equivalent activity for critical appraisal of research papers; half of the EBM teachers used it frequently to always (51% and 50% respectively).

BARRIERS

In Figure 1 median values of barriers for teaching EBM in clinical practice are ranked, and barriers which differ significantly between the countries (overall) are marked. In the text below the level of being a barrier on the Likert scale is expressed as a median with the interquartile range (IQR). Based on median rankings of 120 participants in all countries, severe barriers (median=5) for teaching EBM in clinical practice were overall lack of time for teaching EBM (median 5; IQR 3-6) and lack of time available for trainees to do a literature search (median 5; IQR 3-6). Moderate barriers (median=4) are lack of requirements for EBM, i.e. lack of requirements for EBM skills later in doctors’ career (median 4; IQR 3-6), for EBM skills at exams both at postgraduate (median 4; IQR 3-6) and undergraduate level (median 4; IQR 2-5) and lack of EBM requirements in curricula (median 4; IQR 2-5) and when medical universities are accredited for medical education (median 4; IQR 2-5).
Figure 1: Barriers for teaching EBM in clinical practice (median and IQR): overall and split per country.

<table>
<thead>
<tr>
<th>Barriers</th>
<th>Overall (N=120)</th>
<th>UK (N=29)</th>
<th>Netherlands (N=18)</th>
<th>Germany (N=17)</th>
<th>Switzerland (N=18)</th>
<th>Hungary (N=21)</th>
<th>Poland (N=4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time available for trainees to search the literature*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack requirements EBM later in doctors’ career*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A pyramid hierarchy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack requirements EBM at postgraduate level*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack requirements EBM at undergraduate level*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trainee’s time required reading articles written in English*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack requirements for accreditation medical universities*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of EBM requirements in curricula*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack knowledge and skills to determine applicability to the patient*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack knowledge and skills in defining a relevant search strategy*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of knowledge and skills in critically appraising the evidence*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived lack improvement following EBM, perceived by peers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of assistance in finding evidence*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of hard evidence in discipline*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack knowledge and skills in technical terms used in EBM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack knowledge and skills selecting and using databases*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of trainee’s level of understanding articles written in English*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of teacher’s time required reading articles written in English*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack improvement following EBM, perceived by head department</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack improvement following EBM, perceived by dean*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative attitude trainees towards accepting EBM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of knowledge and skills determining relevant question(s)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack improvement following EBM, perceived by trainees</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Too much evidence without appropriate summaries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flat hierarchy: all clinicians are able to influence practice</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low availability and access to relevant databases*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Barriers marked with an asterix (*) differ significantly between countries (p<0.05)
**Figure 2:** Ranking of the main barriers (with a median ≥3) for the United Kingdom, Hungary, the Netherlands, Switzerland and Poland.

<table>
<thead>
<tr>
<th>Highest median rank</th>
<th>United Kingdom (N=29)</th>
<th>Hungary (N=21)</th>
<th>The Netherlands (N=18)</th>
<th>Germany (N=17)</th>
<th>Switzerland (N=18)</th>
<th>Poland (N=4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of time (4)</td>
<td>Lack of EBM requirements (6)</td>
<td>Lack of time (5)</td>
<td>Lack of EBM requirements (3)</td>
<td>Lack of EBM requirements (6)</td>
<td>Lack of EBM requirements (6)</td>
<td></td>
</tr>
<tr>
<td>Perceived lack of improvement by peers (3)</td>
<td>Lack of time (5)</td>
<td>Lack of time (3)</td>
<td>Lack of EBM requirements (4)</td>
<td>Lack of time (5)</td>
<td>Trainees level understanding English articles (6)</td>
<td></td>
</tr>
<tr>
<td>Lack of knowledge and skills trainers (3)</td>
<td>Lack of EBM requirements (6)</td>
<td>Perceived lack of improvement by peers (3)</td>
<td>Trainees time required reading English articles (4)</td>
<td>Lack of knowledge and skills trainers (4)</td>
<td>Trainees time required reading English articles (6)</td>
<td></td>
</tr>
<tr>
<td>Lack of EBM requirements (3)</td>
<td>Perceived lack of improvement by peers (4)</td>
<td>Lack of knowledge and skills trainers (3)</td>
<td>Trainees level understanding English articles (4)</td>
<td>Trainees level understanding English articles (3)</td>
<td>Lack of time (5)</td>
<td></td>
</tr>
<tr>
<td>Lack of knowledge and skills trainers (4)</td>
<td>Lack of knowledge and skills trainers (4)</td>
<td>Perceived lack of improvement by peers (3)</td>
<td>Lack of assistance in finding evidence (5)</td>
<td>Teachers level of understanding articles written in English (5)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of hard evidence in discipline (4)</td>
<td>Perceived lack of improvement by peers (3)</td>
<td>Lack of assistance in finding evidence (3)</td>
<td>Perceived lack of improvement perceived by trainees (3)</td>
<td>Lack of knowledge and skills trainers (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trainees level understanding articles written in English (4)</td>
<td>Lack of assistance in finding evidence (3)</td>
<td>Perceived lack of improvement perceived by trainees (3)</td>
<td>Lack of hard evidence in discipline (3)</td>
<td>Perceived lack improvement by peers (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Perceived lack of improvement by head department/dean (4)</td>
<td>Lack of hard evidence in discipline (3)</td>
<td>Perceived lack improvement by head department/dean (3)</td>
<td>Perceived lack of improvement perceived by trainees (3)</td>
<td>Lack of evidence summaries (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of evidence summaries (3)</td>
<td>Perceived lack improvement by head department/dean (3)</td>
<td>Lack of evidence summaries (3)</td>
<td>Perceived lack of improvement perceived by trainees (3)</td>
<td>Lack of evidence summaries (3)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative attitude trainees towards EBM (3)</td>
<td>Negative attitude trainees towards EBM (3)</td>
<td>Lack of hard evidence in discipline (3)</td>
<td>Lack of access to internet in outpatient department (3)</td>
<td>Lack of access to relevant databases (3)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
A traditional pyramid hierarchy of junior and senior clinicians and authority of clinical/managerial leadership are, if used, also perceived as major barriers for teaching EBM in practice (median 4; IQR 2-5).

In Figure 2 the ranking of barriers being at least a mild barrier are shown for the United Kingdom, Hungary, the Netherlands, Germany, Switzerland and Poland. For clarity, we grouped the barriers.

Figure 2 shows that although lack of time and lack of EBM requirements are the main barriers in all countries, the trainee’s level of understanding English articles and time required reading English articles are a very severe barrier in Poland (median 6; IQR 5-6), moderate barrier in Germany (median 4; IQR 2-6), Hungary (median 4; IQR 2-6) and Switzerland (median 4; IQR 3-5) and hardly a barrier in the Netherlands (median 2; IQR 1-2). Perceived lack of improvement of patient outcomes following EBM teaching and perceived lack of improvement by peers were considered moderate barriers in Hungary (median 4; IQR 3-5) and mild barriers in Germany (median 3; IQR 2-5) Poland, Switzerland, the Netherlands and the United Kingdom (median 3; IQR 2-4). Additional barriers mentioned by the teachers were; EBM is not used as a tool in health insurance reimbursement policy decisions, there is no trust in EBM models, EBM is not a priority in some organizations and the lack of availability of training rooms.

The total sum of the medians of the barriers was lowest in the Netherlands (total of medians 77) and the United Kingdom (total of medians 84), while rankings were highest in Hungary (total of medians 126) and Germany (total of medians 128) (ANOVA p <0.001).

FACILITATORS
In figure 3 median values of facilitators for teaching EBM in clinical practice are ranked, and facilitators which differ significantly between the countries (overall) are marked. Computer access in clinics and wards is an essential facilitator for teaching EBM and ranked highest in all countries (median 7; IQR 6-7). Improved access to relevant databases or journals (median 6; IQR 6-7), regular teaching activities in the trainers hospital or department (median 6; IQR 5-7), need for EBM skills in quality improvement projects (median 6; IQR 5-7), courses in EBM (median 6; IQR 5-7), requirements for EBM skills when medical universities are accredited for medical education (median 6; IQR 5-6), requirements for EBM skills at exams at postgraduate level (median 6; IQR 5-6), need for EBM skills in policy
### Figure 3: Facilitators for teaching EBM in clinical practice (median and IQR): overall and split per country.

<table>
<thead>
<tr>
<th>Facilitator</th>
<th>Overall (N=120)</th>
<th>UK (N=29)</th>
<th>Netherlands (N=18)</th>
<th>Germany (N=17)</th>
<th>Switzerland (N=18)</th>
<th>Hungary (N=21)</th>
<th>Poland (N=4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Computer access in clinics or wards</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>Improved access to relevant databases/journals*</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>A mentor guiding the teacher on how to teach EBM</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>Regular teaching activities in your hospital/department</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>Courses in EBM</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>Need for EBM skills in quality improvements projects</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>Evidence of improvement following teaching EBM</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>Increased availability or evidence summaries</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>More time allocated for learning EBM for trainee</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>Need for EBM skills in developing guidelines</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>Need for EBM skills in policy decisions</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>EBM requirements at exams at postgraduate level</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>EBM requirements for accreditation for medical education</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>A qualification for EBM teachers</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>More time allocated for teaching EBM for trainer</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>Expectation by the training place to have skills in EBM</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>Availability of clinical librarian(s) with knowledge of EBM*</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>Requirements for EBM skills at exams at undergraduate level</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>E-course on how to teach EBM in native language*</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>Handbook on how to teach EBM in native language*</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>E-course on how to teach EBM in English</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>Handbook on how to teach EBM in English*</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
<tr>
<td>Financial incentives for teaching EBM</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
<td>![Graph]</td>
</tr>
</tbody>
</table>

*Facilitators marked with an asterix (*) differ significantly between countries (P<0.05)
decisions (median 6, IQR 5-6), need for EBM skills in developing guidelines (median 6; IQR 5-6), a mentor guiding the teacher on how to teach EBM (median 6; IQR 4-6), increased availability of evidence summaries (median 6; IQR 5-6), evidence of improvement of clinical practice following teaching EBM (median 6; IQR 5-6), a qualification for EBM teachers (median 6; IQR 4-6) and more time allocated for learning EBM for the trainee (median 6; IQR 5-6) are also very important facilitators for successful implementation of teaching EBM in clinical practice.

Important facilitators were more time allocated for teaching EBM for the trainer (median 5; IQR 5-6), requirements for EBM skills at exams at undergraduate level (median 5; IQR 4-6), expectation by the training place to have skills in providing evidence-based health care (median 5; IQR 5-6) and availability of clinical librarians with knowledge of EBM (median 5; IQR 4-6). Moderately important facilitators were an e-course or handbook on how to teach EBM in native language (median 4; IQR 3-5, for both e-course and handbook) or English (median 4; 3-6, for both e-course and handbook) or financial incentives for teaching EBM (median 4; IQR 3-5).

Moreover, in Hungary and Germany a handbook or e-learning course in the native language scored both significantly higher than an handbook or e-learning course in English (Related Samples Wilcoxon Signed Rank Test p < 0.01). For Switzerland and the Netherlands an English or native tongue course did not score differently (p > 0.20).

Additional facilitators mentioned by the teachers were; summary cards by which you can teach trainees how to critically read an article, an audit of outcomes in the clinical setting, a good team relationship, an EBM mission statement of the department, a few ‘problem-owners’ who are responsible for EBM, advice from experienced clinical epidemiologists, a fee for teaching, free accessibility of the Cochrane Library online, good English-German online dictionaries, international conferences, journals with review articles, organisational commitment, patient empowerment, senior clinicians as role models, technical equipment of course rooms: easy access to the WLAN, trials and reviews actually running in the department, a culture of EBM in the clinic with support from chief of department, a course on what, where and how to search.

Overall facilitators were ranked lowest in Germany (total medians 107) and the Netherlands (total of medians 111) and highest in Switzerland (125) and Hungary (123) – ANOVA p <0,001).
DISCUSSION

In this predominantly European survey, we found that important barriers for teaching EBM in clinical practice were the lack of teaching time in busy practice, lack of curriculum requirements for teaching EBM in clinical practice and the hierarchical nature of the medical profession, which inhibited teaching in light of perceived threat of criticism of seniors. Computer access in clinics and wards was seen as an essential facilitator for teaching EBM on-the-job and ranked highest in all countries. Improved access to relevant databases or journals, regular teaching activities in the hospital or department and formal requirements for EBM skills are very important facilitators for EBM teaching.

Many barriers and facilitators were common in all countries, but there also seem to be differences between countries in the perception of being a barrier or facilitator, e.g. concerning lack of understanding and time required for reading English language articles. We found that overall barriers for teaching EBM were ranked lowest in the Netherlands and the United Kingdom and highest in Poland and Hungary: this might be because in Poland and Hungary the barriers for teaching EBM are truly bigger, or they are perceived higher due to different levels of perception or expectations in the various countries. In any case this study shows a West-to-East gradient of these perceived barriers, which might be explained by historical, cultural, societal, educational and economic and health service differences between these countries.

A study by Matsui et al. in Japan found that lack of English proficiency is the main barrier for teaching EBM, followed by the lack of time. Letelier et al. concluded that language barriers should be taken into account when teaching EBM to Spanish-speaking physicians. Melnyk et al. found resources, including time and money and traditional clinical mindsets/attitudes as main barriers for teaching EBM in nurse practitioner curricula. According to these participants main facilitators were teamwork and mentoring. Meats et al point out that EBM undergraduate teaching is restricted by lack of curriculum time, trained tutors and teaching materials. Davis et al tried to improve evidence based continuing medical education (CME) by an evidence based CME credit designation, but found time constraints and limited understanding of the approval process to be barriers. Similar to this study, those studies also point to difficulties with language and time. Next to their findings we identified other barriers for teaching EBM in practice like lack of
requirements for EBM in curricula in Europe.

A strong point of our study is that, to our knowledge, no previous study focused systematically on a comparison and differences between countries. As clinical teachers ranged from different countries in Europe, we can assume that barriers found in all partner countries will be barriers in practice irrespective of the local specifics of any individual hospital, corporate culture, language or societal values.

There are also some limitations which require a remark. The teachers who filled out the questionnaire might not be representative for all clinical teachers in a country, as most of them filled out the questionnaire as part of their voluntary participation in the piloting of a Training the Trainer EBM course of the EU-EBM project. These participants are inherently more interested in EBM and might see more opportunities or on the contrary more barriers for teaching EBM in clinical practice, which made them decide to take part in the course.

The relatively small number of participants per country makes it difficult to draw firm conclusions about (the magnitude of) differences between countries. This study was exploratory. To be able to adjust for potential confounders, which might mask or introduce differences between countries a larger, representative sample should be used.

To our knowledge, there was no validated questionnaire available which could be used to identify and compare barriers and facilitators for teaching EBM in clinical practice. We therefore constructed a questionnaire ourselves, based on literature review and input from experts in EBM teaching.

Many barriers restrain implementation of on-the-job training in practice, which might implicate that improvements in knowledge and skills of clinical teachers do not automatically imply its teaching in practice. Some barriers can be tackled jointly on a European level, such as requirements for EBM can be laid down in curricula. Other barriers will need to be dealt with on a local level, e.g. translation of materials in the native language or by making good (online) dictionaries and translation programs available.

ACKNOWLEDGEMENTS

We thank all participants for filling out the questionnaire. This study is funded by the EU Leonardo da Vinci program.
REFERENCES

CHAPTER 10

TOOLS TO ASSESS EVIDENCE-BASED PRACTICE BEHAVIOUR AMONG HEALTHCARE PROFESSIONALS
A SYSTEMATIC REVIEW

Katrien Oude Rengerink, Sandra E. Zwolsman, Dirk T. Ubbink, Ben Willem J. Mol, Nynke van Dijk, Hester Vermeulen

Evidence-based Medicine, 2013;18:129-38
ABSTRACT

**Objective** - To identify and compare tools to assess Evidence-Based Practice (EBP) behaviour among healthcare professionals.

**Design** - Systematic review.

**Data sources** - MEDLINE, EMBASE, Cochrane Library, PsychInfo and CINAHL up to July 2011.

**Study selection** - Titles, abstracts and eligible full text articles were screened by two reviewers independently.

**Data extraction** - Relevant data were extracted by one reviewer and checked by a second reviewer. Eligibility criteria for selecting studies: original studies among all healthcare professionals that described the development or use of EBP behaviour assessment tools.

**Results** - Of 19,310 identified articles, 172 studies were included. We identified 117 questionnaires, 10 interviews or focus groups, nine observational studies, 27 chart evaluations and nine studies used a combination of methods. Psychometric properties of the questionnaires used were reported in about half of the studies, in seven studies that assess a single EBM step and in six studies that assess a combination of EBM steps. One of these assessed all five steps of EBP.

**Conclusions** - Valid and reliable EBP behaviour assessment tools are available. However, only one questionnaire validly and reliably assessed all five EBP steps, covering the entire concept of EBP.
INTRODUCTION

Healthcare decisions should preferably be based on high-quality research evidence such as clinical guidelines, systematic reviews or randomised clinical trials.\(^1\) Ironically, healthcare professionals often fail to implement clinical procedures that have established efficacy or fail to discard proven ineffective procedures.\(^2\) A study in the USA suggests that approximately 30% of patients do not receive care in accordance with the latest scientific evidence and approximately 25% of patients receive unnecessary or potentially harmful care.\(^3\) The gap between evidence and practice still exists.

To overcome the gap between best practice and actual care, professional organisations worldwide encourage Evidence-Based Practice (EBP).\(^4\) The five steps of EBP—Ask, Access, Appraise, Apply and Assess—equip healthcare professionals with the necessary steps to successfully integrate evidence from research with their clinical decision-making.\(^5\) Competency in EBP has become a prerequisite for (re)certification of healthcare professionals.\(^6,7\) Although EBP has more and more become the standard of care, there are still barriers to overcome that refrain health professionals from teaching or working evidence based, like a lack of time to read evidence, lack of facilities or resources, lack of requirements for EBM, lack of EBM skills, a pyramid hierarchy in healthcare management structure discouraging EBM, and barriers related to the available evidence.\(^8,9\)

To be able to assess whether healthcare professionals and healthcare organisations actually work evidence based, a valid and reliable method for the assessment of EBP behaviour in clinical practice is needed. EBP behaviour can be assessed by considering if, and at what level, individual healthcare professionals use the five EBP steps in daily practice.\(^7\) Alternatively, the application of evidence-based clinical manoeuvres could be assessed.\(^10\) The optimal method for evaluation of EBP behaviour is unclear. Shaneyfelt et al\(^11\) reviewed tools that evaluate EBP, but they focused on evaluating the effect of teaching EBP. To evaluate EBP teaching, most often knowledge and skills were assessed, rather than impact on daily clinical practice.\(^11\) Their review showed that the Fresno Test and the Berlin Questionnaire are valid and reliable for assessing knowledge and skills of individual trainees.\(^12,13\) But, as improvement in knowledge and skills does not automatically lead to an improvement of behaviour in practice, it is important to measure actual EBP behaviour as well.\(^6\) Shaneyfelt et al identified four valid EBP behavior instruments using objective outcome measures, but these instruments did not have the ability to document the EBP
behaviour of individual professionals.\textsuperscript{5, 11–13}
An overview of existing EBP behaviour assessment tools could help to determine the optimal assessment method. Therefore, we systematically reviewed the validity, reliability and feasibility of all existing methods to assess EBP behaviour of healthcare professionals.

METHODS

This review was performed and described according to the PRISMA statement, using a prespecified protocol.\textsuperscript{14}

ELIGIBILITY CRITERIA

We included original studies among all healthcare professionals (i.e., physicians, dentists, nurses and other allied healthcare professionals such as physiotherapists, speech-language therapists, occupational therapists and dental hygienists) that described the development or use of EBP behaviour assessment tools.

We excluded studies about adherence to guidelines and studies about evidence-based care or quality indicators regarding one particular disease, since these tools address specific behaviour regarding the guideline or disease evaluated and outcomes of these studies would likely be hard to extrapolate to other (general) settings. Furthermore we excluded studies about the evaluation of preclinical students, as they are not working in practice yet. To optimise applicability of the results we excluded randomised controlled trials (RCTs) that evaluate strategies for improving EBP behaviour, because the evaluation used to assess the strategies may not be feasible outside the trial. Proceedings of conferences were not included as they contained too little information about the assessment methods used.

INFORMATION SOURCES AND SEARCH

A search for eligible studies was performed in MEDLINE (Pubmed), EMBASE (Ovid), the Cochrane Library, CINAHL (EBSCOhost) and PsychINFO (EBSCOhost) without any restrictions to language from the earliest available date until July 2011. The search terms are listed in online supplementary appendix 2. We did not restrict our search to studies evaluating psychometric properties of instruments, as we expected not to find many validated instruments while this restriction would result in the possible exclusion of relevant, but not yet validated instruments.
STUDY SELECTION
In pairs, two reviewers independently reviewed the titles and abstracts of the retrieved studies for eligibility. Of the selected studies, the full articles were appraised by two reviewers to determine eligibility for inclusion. In the case of persisting disagreement during any step in the review process, a third reviewer was consulted.

DATA COLLECTION PROCESS AND DATA ITEMS
A structured data extraction form (see online supplementary appendix 3) was used to collect the following relevant data from the included studies:

1. Characteristics of the participating healthcare professionals: number, discipline, training level;
2. Description and development of the EBP behavior assessment tool as described by the study authors;
3. The classification of the tool regarding the five EBP steps (Ask, Access, Appraise, Apply and Assess) or patient outcomes;
4. Psychometric properties of the tool as described in the original study with a notification whether reliability, validity and responsiveness had been tested;
5. Any previously developed methods on which the behaviour assessment tool had been based;
6. Description of whether the assessment is subjective (i.e., self-reported measures) or objective (formal assessment or actual observations in practice).

We also extracted the country in which the instrument was developed. Countries were plotted on a world map using R 2.14.0, package Maptools (version 0.8-21). The data of the included studies were extracted by one reviewer and confirmed or corrected by a second reviewer. Disagreements during data extraction were resolved during a consensus meeting. A third reviewer was consulted in case of persisting disagreement.

RISK OF BIAS ASSESSMENT OF INDIVIDUAL STUDIES
We did not perform a quality assessment of the included studies, as a well-performed study does not guarantee good quality of the assessment tool used and vice versa.
RESULTS

STUDY SELECTION
The search resulted in 19310 titles and abstracts. Of 326 abstracts the full articles were retrieved, of which 172 studies met all inclusion criteria, representing 156 different behaviour assessment tools. A flowchart of study inclusion is shown in figure 1.

Figure 1: Flow chart of the search strategy and number of articles retrieved.\textsuperscript{14}

<table>
<thead>
<tr>
<th>Identification</th>
<th>Nr of records identified through database searching 28,080</th>
</tr>
</thead>
<tbody>
<tr>
<td>Screening</td>
<td>Nr of records after duplicates removed 19,310</td>
</tr>
<tr>
<td></td>
<td>Nr of records screened 19,310</td>
</tr>
<tr>
<td></td>
<td>Nr of records excluded 18,984</td>
</tr>
<tr>
<td>Eligibility</td>
<td>Nr of full text articles assessed for eligibility 326</td>
</tr>
<tr>
<td></td>
<td>Nr of full text articles excluded, reasons:</td>
</tr>
<tr>
<td></td>
<td>• 10 full text not retrieved</td>
</tr>
<tr>
<td></td>
<td>• 9 duplicate</td>
</tr>
<tr>
<td></td>
<td>• 59 no original peer reviewed study</td>
</tr>
<tr>
<td></td>
<td>• 68 no clear description of method measuring EBM</td>
</tr>
<tr>
<td></td>
<td>behaviour of health professional in practice</td>
</tr>
<tr>
<td></td>
<td>• 1 randomized trial</td>
</tr>
<tr>
<td></td>
<td>• 7 adherence to 1 organ, guideline, disease</td>
</tr>
<tr>
<td>Included</td>
<td>Nr of studies included in qualitative synthesis 172</td>
</tr>
</tbody>
</table>

STUDY CHARACTERISTICS
These studies were retrieved from nearly every continent (figure 2). Tools to measure EBP behaviour were first developed in 1992. In online supplementary appendix 1 an overview is presented of all tools in which a description of validity or reliability was included and in which validity and reliability were tested and/or established.

The remaining tools were used to assess other healthcare professionals: dentists, dental hygienists, pharmacists, physiotherapists, mental health practitioners, public health workers, speech-language pathologists, social health workers, occupational therapists, ambulance officers, dieticians or a mix of professionals.

Mostly, the EBP behaviour assessment was part of a broader assessment tool, for instance a tool that also assesses EBM knowledge, skills or attitude. The methods used to evaluate
behaviour were questionnaires, interviews and focus groups, observations or registration of healthcare professionals, evaluation of charts or a combination of methods.

QUESTIONNAIRES
Of the EBP behaviour assessment tools described in the various studies, 117 concerned questionnaires. Table 1 shows 42 EBM questionnaires for which validity and/or reliability have been tested and described—not necessarily confirmed. Two of these questionnaires concerned translations of existing tools. A list of questionnaires or tools in which validity or reliability have not been tested (if applicable) is available on request. Of these 42 studies, 16 studies assessed only one step of EBP: two assessed step 1 (Ask), nine step 2 (Access) and five step 4 (Apply). In 26 questionnaires multiple EBP steps were assessed, mostly a combination of steps 2 and 4 (Access and Apply). Most questionnaires were to be completed by the respondent. Some were structured telephone surveys conducted and filled in by a researcher. In about 60% of the identified questionnaires previously reported questionnaires were used for further development, of which the questionnaire by McColl et al was most frequently used. The psychometric properties of the questionnaires were reported in only about half of the studies: 28 studies tested validity and 20 tested reliability, of which 11 tested both validity and reliability.
Table 1: Overview of EBM questionnaires of which validity and/or reliability have been tested (not necessarily confirmed).

<table>
<thead>
<tr>
<th>EBM step</th>
<th>Author</th>
<th>Healthcare professional</th>
<th>Description of the instrument</th>
<th>Based on</th>
<th>Reliable</th>
<th>Valid</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Carter</td>
<td>433 physical therapists</td>
<td>Reading the literature without a question.</td>
<td>Eckerling 1988</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Karlsson</td>
<td>425 occupational therapists</td>
<td>Read occupational therapy research literature to update knowledge.</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>2</td>
<td>Hendricson</td>
<td>16 dental faculty</td>
<td>KACE: Evidence Based Practice Knowledge, Attitudes, Access, and Confidence Evaluation. Accessing evidence: frequency of accessing the Cochrane Library.</td>
<td>Taylor 2001; Fritsche 2002; Johnson 2003; Bradley 2004;</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Ross</td>
<td>128 anaesthesia nurses</td>
<td>Information Literacy for Evidence-based Nursing Practice: information seeking and resource use.</td>
<td>Pravikoff 2005</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chiu</td>
<td>1156 MD and nurses  457 MD</td>
<td>Frequency database access to search for medical information.</td>
<td>Yu 2001; Wong 2005; Bennett 2005</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Thiel</td>
<td>121 registered nurses</td>
<td>Frequency of searching information, using modified Information Literacy for Evidence-based Nursing Practice.</td>
<td>Pravikoff 2005; Pierce 2000; Tanmer 2000</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Shirkhedkar</td>
<td>106 medical officers</td>
<td>Accessibility, frequency of use and preference for electronic databases.</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Phua</td>
<td>154 MDs various disciplines</td>
<td>Time spent on UpToDate and other information resources.</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Carney</td>
<td>129 community clinical teachers</td>
<td>Use of information sources for patient care decisions: the Internet, Medline, MD consult.</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Jette</td>
<td>488 physiotherapists</td>
<td>Frequency of reading articles, use of databases, literature and access to practice guidelines.</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Philibert</td>
<td>328 occupational therapists</td>
<td>Reading journals and sources used.</td>
<td>Kirk 1976; McKee 1987</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>4</td>
<td>Forsman</td>
<td>2621 nurses</td>
<td>Use of research: direct research use, indirect research use and persuasive research use (questions unclear)</td>
<td>Estabrooks 1997, 1999, Formans 2009 (translation Estabrooks)</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Cobban</td>
<td>161 dental hygienists</td>
<td>Research utilization.</td>
<td>Estabrooks 1998</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Olade</td>
<td>106 nurses</td>
<td>Utilization of nursing research and type of research.</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Veeramah</td>
<td>173 nurses and midwives</td>
<td>Frequency of using research findings in practice.</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Lacey</td>
<td>20 nurses</td>
<td>Research utilization; reading more than or equal to 11 English articles in past 6 months.</td>
<td>Champion 1989</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>1,2</td>
<td>Chernick</td>
<td>56 MDs: paediatric interns and residents; EBP experts</td>
<td>Self-reported practice of EBP: frequency of searching articles to answer clinical question, generation of clinical questions applicable patients’ diagnostic or therapeutic plan.</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>EBM step</td>
<td>Author</td>
<td>Healthcare professional</td>
<td>Description of the instrument</td>
<td>Based on</td>
<td>Reliable</td>
<td>Valid</td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
<td>-------------------------</td>
<td>-------------------------------</td>
<td>----------</td>
<td>----------</td>
<td>-------</td>
</tr>
<tr>
<td>2,4</td>
<td>Fillipini(^{34})</td>
<td>449 of 923 nurses</td>
<td>Frequency of reading guidelines and scientific journals; having modified practice last year; frequency of using EBP.</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Profetto-McGrath(^{35})</td>
<td>94 nurses</td>
<td>Cross sectional telephone survey: frequency of accessing written or people-based evidence; and how evidence is used.</td>
<td>Profetto-McGrath 2007</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abeysana(^{36})</td>
<td>315 MD</td>
<td>Number of articles read per month and use of EBM in the management of patients.</td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Guindon(^{37})</td>
<td>1499 GPs, specialists, nurses, health workers</td>
<td>Access to evidence and use of sources; use of evidence and change in practice attributed to particular sources of research evidence.</td>
<td>Landry 2003; McColl 1998; Calverton 2004; Page 2003; Prescott 1997; Wilson 2001 and 2003; WHO 2002 and 2004; Canadian Health Services Research Foundation 2001</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>McEvoy(^{38})</td>
<td>105 (mixed)</td>
<td>Tracked down relevant evidence once a question formulated; integrated research evidence with expertise;</td>
<td>Kamwendo 2001; Green 2002; Jette 2003; Iles 2006; Bridges 2007</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Adams(^{39,40})</td>
<td>386 school nurses</td>
<td>SN EBP questionnaire: which sources used when information needed.</td>
<td>Titler 1999; Rodgers 2003</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ahmad(^{41})</td>
<td>332 primary care MDs</td>
<td>Self-reported amount of EBP; frequency of use of different sources.</td>
<td>McColl 1998; McKenna 2004</td>
<td></td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Kahveci(^{42})</td>
<td>375 primary care physicians</td>
<td>Use of resources and percentage of practice evidence based.</td>
<td>McColl 1998; Al-Ansary 2002; Al-Almaie 2004</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Melnyk(^{43})</td>
<td>160 nurses</td>
<td>Use of Cochrane Database; amount of practice evidence-based.</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Tranmer(^{44})</td>
<td>235 nurses</td>
<td>Research Utilization Questionnaire</td>
<td>Champion 1989</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Johansson(^{45})</td>
<td>99 head nurses</td>
<td>Searching and EviPraQ about evaluation of care.</td>
<td>Champion 1989</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2,5</td>
<td>Wallin(^{46})</td>
<td>119 nurses</td>
<td>Seeking new research literature and research use in daily practice.</td>
<td>Humphris 1999; Champion 1989; Pettengil 1994</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>4,5</td>
<td>Brown(^{47})</td>
<td>696 paediatric occupational therapists</td>
<td>KAP survey: research knowledge, attitudes and practices</td>
<td>van Mullem 1996</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>1,2,3</td>
<td>Amin(^{48})</td>
<td>348 MDs</td>
<td>How frequently do you encounter information gaps; what proportion of practice is evidence based.</td>
<td>McColl 1998; Olantunbosun 1998; Fritsche 2002.</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thomas(^{49})</td>
<td>59 paediatric dieticians</td>
<td>How often do you encounter a knowledge gap and what domain; literature search yes/no; information sources used most often; criteria taken into consideration when critically appraising paper.</td>
<td></td>
<td></td>
<td>x</td>
</tr>
</tbody>
</table>

* SN EBP, School Nurse Evidence Based Practice Questionnaire
The remaining studies did not report any psychometric properties, although some studies mentioned that their questionnaire was based on literature—without any further specifications—or a pilot study was performed.

For the assessment of EBP behaviour by means of questionnaire administration, several tools were identified that were shown to be valid and reliable. For the assessment of behaviour of single EBP steps, the valid and reliable tools are: step 2: Hendricson, Chiu, Philibert and Jette and step 4: Cobban and Veeramah.

For the assessment of multiple EBP steps, the following tools were found to be valid and/or reliable: steps 1 and 2: Chernick, steps 2 and 4: Fillipini, Kahveci, Guindon and McEvoy.

To assess all five EBP steps, the questionnaires by Scott and Wallen were tested for reliability, but not for validity. The questionnaire of Bostrom showed adequate validity, but was not tested for reliability. Six items measuring the respondents' extent of applying the components of EBP, on a four-point response format (1=to a very low extent, 4=to a

<table>
<thead>
<tr>
<th>EBM step</th>
<th>Author</th>
<th>Healthcare professional</th>
<th>Description of the instrument</th>
<th>Based on</th>
<th>Reliable</th>
<th>Valid</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,2,4</td>
<td>Parahoo</td>
<td>87 nurses</td>
<td>Frequency of reading research studies, info sources, implementation of research findings.</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Poolman</td>
<td>367 orthopaedic surgeons</td>
<td>Use of EBM in clinical decision making.</td>
<td>McColl 1998</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>2,3,4</td>
<td>Prior</td>
<td>55 nurses</td>
<td>Three subscales, 24 items: practice EBP (never-frequently), tracked down evidence; critically appraised; integrated evidence.</td>
<td>Upton 2006</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td>1,2,3,4,5</td>
<td>Bostrom</td>
<td>1256 nurses</td>
<td>Use of EBM: formulate questions; seek relevant knowledge, critical appraisal, implementation, evaluation.</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Scott</td>
<td>111 MDs</td>
<td>Number of questions formulated per week; number of searches for evidence performed per week; databases, topics, critical appraisal, efficiency of search in yielding information from patient management; applying in practice, confidence and change in practice; questionnaire evaluating practice.</td>
<td></td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Romani</td>
<td>459 nurses</td>
<td>Frequency of use of EBM applied in care for individual patients.</td>
<td>Upton 2006 (translation)</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Wallen</td>
<td>159 nurses</td>
<td>Questionnaire: EBP implementation scale (EBPI); frequency of EBP implementation behaviour over the past 8 weeks (e.g. used evidence to change clinical practice; critically appraised evidence)</td>
<td>Melnyk 2003</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Estrada</td>
<td>594 nurses, acute care</td>
<td>Unclear, the 18 item EBP implementation scale of Melnyk was used.</td>
<td>Melnyk 2008</td>
<td>x</td>
<td></td>
</tr>
</tbody>
</table>
very high extent): (1) formulate questions, (2) seek relevant knowledge using databases, (3) seek relevant knowledge using other information sources, (4) critically appraise and compile best knowledge, (5) participate in implementing research knowledge in practice and (6) participate in evaluating practice based on research knowledge.

Romani\textsuperscript{55} translated the questionnaire developed by Upton, which assesses all five EBP steps and has good psychometric properties—as tested by Upton (but this is not further described). This questionnaire consists of three parts, of which the first part focuses on EBM behaviour. It is a self-reported questionnaire that asks how often in the last year: (1) A question was formulated to fill a knowledge gap, (2) How often was searched for evidence, (3) How often critical appraisal was applied, (4) How often the literature that has been found was integrated with own experience and knowledge, (5) The results were evaluated in practice and (6) The knowledge was shared with colleagues.

INTERVIEWS AND FOCUS GROUPS
Interviews or focus groups were used in five studies,\textsuperscript{59–63} a short description of these studies can be found in the online supplementary appendix. A variety of EBP steps was measured in the interviews, most of the interviews included step 2 or 4. Two studies described validation of their tool: Bogdan Lovis\textsuperscript{59} and Rolfe.\textsuperscript{60}

OBSERVATIONS OR REGISTRATION OF HEALTHCARE PROFESSIONALS
The EBP behaviour of healthcare professionals was directly observed in nine studies,\textsuperscript{64–72} an overview of these studies can be found in table 2. Observation commonly entails the mean logins to the library or database,\textsuperscript{64} the use of articles, or (a change in) decision-making as a result of searches,\textsuperscript{67} information seeking behaviour,\textsuperscript{65,68} a portfolio\textsuperscript{65} or the evidence base for expert opinions.\textsuperscript{70} None of the identified tools mentioned aspects of validity and reliability; however, it can be discussed whether it is necessary to do so when using observations. Tilburt et al\textsuperscript{67} did test interobserver agreement regarding the observation of EBP steps 1-3.

EVALUATION OF CHARTS
Evaluation of charts was performed in 27 studies to assess the evidence base underlying diagnosis-intervention pairs, surgeries or interventions during pregnancies,\textsuperscript{73–99} an overview of these studies can be found in the online supplementary appendix. In 12 of these evaluations the level of evidence had been based on the criteria according to Ellis
et al\textsuperscript{10}: evidence from RCTs (level 1), convincing non-experimental evidence (level 2), interventions without substantial evidence (level 3). The classification by Kingston et al\textsuperscript{90} was also frequently used. Three studies tested the validity of their tools: Kenny\textsuperscript{81} (also tested reliability); Nordin-Johansson\textsuperscript{96} and Straus.\textsuperscript{76}

Table 2: Overview of real time observation or registration of healthcare professionals EBM behavior in practice

<table>
<thead>
<tr>
<th>EBM step</th>
<th>Author</th>
<th>Healthcare professional</th>
<th>Description of the instrument</th>
<th>Based on</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>Haynes\textsuperscript{64}</td>
<td>203 MDs</td>
<td>Mean registered logins to the library per user: use of PLUS (peer-selected articles): percentage users and minutes/month.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Van Duppen\textsuperscript{65}</td>
<td>2920 doctor patient contacts, 365 searches, 5 GPs</td>
<td>Online on the spot searching. Number of searches; number of databases consulted; answer found, change of decision.</td>
<td>Shaughnessy 1994; Slawson 2005</td>
</tr>
<tr>
<td>2,3,4</td>
<td>Fung\textsuperscript{66}</td>
<td>41 residents in obstetrics and gynaecology</td>
<td>Learning portfolio: questions reported and directly forwarded to searching strategy; whether or not a change in practice occurred.</td>
<td></td>
</tr>
<tr>
<td>1,2</td>
<td>McKnight\textsuperscript{67}</td>
<td>6 registered nurses</td>
<td>Participant observation and in-context interviews were used to record in detail fifty hours of the information behaviour of a purposive sample of on-duty critical care nurses on critical care unit.</td>
<td></td>
</tr>
<tr>
<td>1,2</td>
<td>Tilburt\textsuperscript{68}</td>
<td>70 general internal medicine residents</td>
<td>Resident behaviour was observed and resident–attending doctor interactions audio taped looking for themes of information exchange, using the method by Bernard 1995.</td>
<td></td>
</tr>
<tr>
<td>1,2,3</td>
<td>Darst\textsuperscript{69}</td>
<td>10 paediatric cardiologists</td>
<td>Ten paediatric cardiologists recorded every clinically significant decision made during procedures, test interpretation or delivery of inpatient and outpatient care. The basis for each decision was assigned to one of 10 predetermined categories, ranging from arbitrary and anecdotal, to various qualities of published studies, to parental preference and avoiding a lawsuit.</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Linthorst\textsuperscript{70}</td>
<td>25 cases of exotic expert opinion, department of internal medicine</td>
<td>Of 25 exotic expert opinions, based on careful literature review of 8 the statements were evidence based, contradicting literature was found in 13 and no literature in 4.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Qian\textsuperscript{71}</td>
<td>599 women from obstetrics ward</td>
<td>Using Cochrane Library 6 procedures were selected which should be avoided as routine and two that should be encouraged. The procedure rate was determined by exit interviews with women, verified using hospital notes.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Coffey\textsuperscript{72}</td>
<td>22 gerontological nurses</td>
<td>Portfolio: documentation of identification of ways for improvement, documentation of sources of research and evidence for practice.</td>
<td>Fensom 2000</td>
</tr>
</tbody>
</table>
COMBINATION OF MULTIPLE METHODS
A combination of the assessment methods as aforementioned was used in six studies being mostly a combination of interviews with observations or questionnaires.\textsuperscript{100–105} Luker et al\textsuperscript{101} combined observations, individual interviews, group discussions and records. Only Cullen established validity and reliability for steps 2 and 3.\textsuperscript{104}

DISCUSSION
This review identified 172 studies representing 160 different tools that assess EBP behaviour. Most of these subjectively assessed a single step of EBP without establishing psychometric properties. Valid and reliable tools for the assessment of a single EBM step or a combination of EBM steps were identified, but no tool with established validity and reliability assesses all five EBP steps. For one questionnaire measuring all five steps validity has been established.\textsuperscript{53}

Strengths of our review are that it includes assessment tools for all health professionals, reducing the chance of missing a potential important assessment tool that could be extrapolated from one health setting to another. Moreover, it focuses on behaviour in clinical practice, which has a more direct effect on patient outcomes than knowledge and skills.
This review has some points which require a remark. First, the definition of EBP as used worldwide is not always congruent: sometimes EBP is defined as the adherence to guidelines, whereas others regard EBP as the integration of all components of EBP in clinical practice. Second, we could have missed studies during the review process. Since EBP behaviour is a broad concept, we added search terms to our search strategy to decrease the number of retrieved studies. However, as we searched the literature with a broad search strategy revealing over 19,000 hits we think it is unlikely we missed any ideal and frequently used method. Lastly, in this review we did not explore the relation between the use of EBP and the level of knowledge and skills of the professional in the varying healthcare settings. These should, however, also be considered when choosing the optimal method for measuring EBP behaviour. Enacting EBP behavior without the proper knowledge and skills can lead to wrong conclusions.\textsuperscript{106} When the instrument will be used to assess health professionals and/or institutions for their EBP behaviour, this might be combined with a measurement of EBP competency, like knowledge, skills and/or attitudes. However, this review did not focus on these domains.
Tools measuring EBP behaviour of healthcare professionals should assess the use of (one of the) EBP steps in practice, the performance of evidence-based clinical manoeuvres and/or the effect of EBP on patient outcomes. Tilson et al\textsuperscript{7} stated that all five EBP steps a clinician would use in practice need to be considered. The tools identified in this review most frequently appreciate accessing evidence (step 2), which allows only a narrow view on EBP behaviour neglecting the integration of evidence with patient values and clinical expertise.

Because we did not restrict our review to specific healthcare settings, the tools presented here could be applied to a wide range of healthcare professionals. However, each group of healthcare professionals might have specific needs or barriers for EBP. For example, a study shows for non-English speaking nurses understanding English articles can be a large barrier, while for physicians this seems a less important problem.\textsuperscript{108} Therefore, the selection of the optimal or most appropriate tool might be tailored to the group of healthcare professionals.

When choosing the optimal method, the feasibility of the instrument should also be considered: interviews and observations are more time-consuming for the evaluator than questionnaires. However, interviews and observations may give more in-depth information. In this study we did not assess the feasibility of instruments, although little information on this aspect was reported in the included studies. The questionnaire of Boström\textsuperscript{53} showed adequate validity, which seems promising, but was not tested for reliability. However, EBP behaviour is measured using only six items. As the questionnaire is self-reported, this might introduce bias, because of social desirable answers. The questionnaire is constructed to assess the EBP level of Swedish nurses, so its generalizability should be assessed before adopting it in practice.

In conclusion, this review identified tools that validly and reliably assess single steps of EBP behaviour. One tool measures all aspects of EBP behaviour validly, but the reliability of this tool has not been established yet.\textsuperscript{53} For future developmental studies this tool should be evaluated more extensively and/or existing valid and reliable tools could be combined into an instrument that covers all EBP steps. Evaluating EBP behaviour is important to deliver optimal healthcare. For proper evaluation it is necessary to predetermine the EBP behaviour that is to be expected, the aim of the assessment and the context and setting of the healthcare profession to be assessed.
Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/eb-2012-100969).

Appendix 1
Overview of EBM behavior assessment tools of which validity and/or reliability have been tested (extension table 1).

Appendix 2
Search strategy used to retrieve selected articles.

Appendix 3
Data abstraction form.
REFERENCES


This thesis presents a number of research projects centered on ‘evidence-based medicine’ (EBM). In the projects reported here, we focused both on the generation of evidence, through clinical trials, and the integration of evidence from solid research into clinical practice.

This thesis consists of two parts. Part 1 focuses on improving recruitment of the necessary number of patients in clinical trials, as this is a major problem while evaluating the effectiveness of interventions in health care. To improve our understanding of patient recruitment, we tried to identify obstacles and facilitators for successful recruitment. Part 2 of this thesis focuses on improving integration of evidence-based decision making in clinical practice. We identified barriers for EBM teaching in practice, developed an EBM Teach the Teacher course, and evaluated methods to assess EBM in clinical practice.

In Chapter 1 we introduced the rationale for this thesis. ‘Evidence-based medicine’ (EBM), defined in 1992 by Gyatt as “the integration of individual clinical expertise with the best available external evidence from systematic research”, is considered the cornerstone of decision making in current clinical medicine. As the way in which we intuitively derive knowledge from experience can lead to biased decisions, we need valid evidence from high quality clinical trials.¹

PART 1: IMPROVING RECRUITMENT OF PATIENTS IN CLINICAL TRIALS

In Chapter 2 the rationale and design of our project “IMproving PArticipation of patients in Clinical Trials” (IMPACT) were presented. To better understand barriers and facilitators for successful recruitment in trials from different levels, we planned two cohort studies (Chapter 3 and Chapter 4) and a case-control study (Chapter 5).

Chapter 3 depicted factors that influence successful recruitment in trials, based on an analysis from a cohort of 1,129 trials registered in the Netherlands Trial Register with a planned completion date between 2005 and 2010. We analyzed 386 trial questionnaires, reporting on a total recruitment target of 42,412 participants in multiple specialties, both in industry and non-industry sponsored trials. Two thirds (66%) of the trial coordinators stated that recruitment was more difficult than anticipated. In less than half (43%) of the trials 80% or more of the targeted number of patients had been recruited at the planned completion date, while a mere 53% had reached the targeted sample size at the
completion date. Only 52% of the trials had been reported in the peer reviewed literature within two years after the completion date. Unfortunately we failed to identify a consistent pattern of characteristics associated with slow recruitment.

Chapter 4 described the results of a web-based questionnaire sent to the local researchers of centres that had recruited in total 14,808 patients for 17 multicentre trials performed in the Consortium for Obstetrics and Gynaecology in the Netherlands. We studied whether factors motivating centre decisions about participation in trials, the research orientation of the department, and the (perceived) logistical support by research personnel were associated with recruitment. In univariable analysis, participating in trials because “expecting others to recruit in return for their studies” was associated with higher recruitment scores, while participating “because it is expected from us” was associated with lower scores. Higher recruitment scores were seen in centres regularly initiating research and in tertiary care centres. Weighted recruitment scores were higher in centres in which research staff was more frequently available.

In Chapter 5 we reported findings from 21 semi-structured interviews with pregnant women invited to a randomized trial in obstetrics. We studied two groups: those who participated (n=12) and those who had declined participation (n=9). We interviewed patients from 8 different trials. Contribution to scientific research was for 5 of 12 participants the main motive for participation in the trial, while 5 mentioned to have participated because the intervention was not available outside the trial. Key motives for non-participation (n=9) were a negative association or dislike of the intervention, either because it might do harm (n=6) or for practical reasons (n=3). We identified seven main themes that influenced decisions on participation. We noted that uncertainty about scientific research and/or the intervention was reported to be of considerable importance. Patients’ understanding of scientific research was generally low. A personal, complete and well-timed dialogue may facilitate a balanced decision on trial participation.

Chapter 6 presented a questionnaire survey sent to clinicians in all departments of Obstetrics and Gynecology in the Netherlands. We observed that recruitment for multicenter clinical trials was associated with better knowledge about the trial results; overall, clinicians who had recruited were on average 1.8 times more likely to know the trial results, compared to clinicians who had not recruited (95% confidence interval (CI) 1.7 to 1.9; range between 1.1 and 3.3 across individual studies). Recruitment for a trial was
also significantly associated with implementation of study results, though the effect was small, with a summary relative risk of 1.1 (95% CI 1.02 to 1.19).

PART 2: IMPROVING EVIDENCE-BASED MEDICINE IN CLINICAL PRACTICE

Many courses on evidence-based medicine (EBM) are offered nowadays, but optimally integrating EBM teaching in clinical practice remains a challenge. Teach the Teacher (TTT) EBM courses, aimed at educating clinical EBM teachers, could help to improve teaching quality and to set a standard for EBM teaching.

In the project described in Chapter 7 we evaluated the availability and content of existing TTT EBM courses. Through country specific EBM networks and infrastructure we identified 16 courses that could be labelled as TTT courses: 4 in Hungary, 4 in the Netherlands, 4 in the United Kingdom, and 4 in Germany. Of these, 15 out of 16 were taught by academic specialists and 11 by methodologists or statisticians. Seven of the course organizers highlighted the need for help in providing practical examples of successful techniques for teaching EBM in clinical practice, 7 for help with a curriculum for trainers to train healthcare workers, and 7 help with funding of TTT courses.

Chapter 8 presented an evaluation of a web based educational course for clinical trainers, designed to teach EBM principles in everyday practice with confidence. The course consisted of e-learning modules that defined the learning objectives and incorporated video clips of practical and effective EBM teaching methods, aimed to seize opportunities for teaching in six different clinical settings. In a group of 56 clinical tutors, in different specialties across six European countries, 47 (84%) moderately improved their score on a questionnaire to measure learning achievement after the course, from an average of 69 (SD 10) to 77 points (SD 12) (p<0.0001).

Improving knowledge on how to teach EBM does not automatically translate in more or better EBM teaching in clinical practice. In Chapter 9 we identified and compared barriers for teaching EBM in clinical practice in various European countries, using a questionnaire completed by 120 clinical EBM teachers from 11 countries. On a 7 point Likert scale, anchored at “1: not at all” on the one end and “7: an insurmountable barrier” at the other, “lack of time” received the highest score (median 5). Moderate barriers (median 4) were “lack of requirements for EBM skills” and “a top down health care management structure”.

In Germany, Hungary and Poland reading and understanding articles in English received higher barrier scores than in the other countries.

Self-reported barriers and facilitators can be of help in identifying strategies to improve practice. To measure whether such strategies have an effect on actual behavior, a valid and reliable assessment method is needed. In Chapter 10 we presented a systematic review of original studies that described the development or use of tools to assess Evidence-based Practice behavior in healthcare professionals. We identified 172 such studies: 117 questionnaires, 10 interviews or focus groups, 9 observational studies, and 27 chart evaluations, while 9 other studies used a combination of methods. Psychometric properties of the questionnaires used were reported in about half of the studies. In 7 studies, assessing a single EBM step, and in 6 studies that assess a combination of EBM steps with a questionnaire, validity and reliability of the instrument were both reported. One of these questionnaires assessed all five steps of EBM. We concluded that valid and reliable methods are available, but the specific aims of the assessment should guide the choice of an instrument for assessing EBM behavior in practice.
IMPLICATIONS FOR PRACTICE AND SUGGESTIONS FOR FURTHER RESEARCH

The studies summarized in this thesis and comparable projects by others have shown that recruiting a sufficient sample of patients and improving evidence-based clinical decision making are complex, context dependent processes. In general, we found that patient recruitment and EBM seem to flourish better when sowed in fertile ground: an environment where recruitment for trials and working evidence-based are (culturally) embedded in clinical practice and accepted, or even expected, by opinion leaders and clinical trainers. Yet a fertile ground in itself is not a sufficient condition for clinicians to change from a more authority based culture to a culture where the rationale underlying decisions is critically appraised and openly discussed. Inviting patients to participate in trials presupposes the courage to admit that the medical community currently does not know what is best, while EBM directly or indirectly challenges the traditional authority of trainers and clinicians. A large number of implementation strategies have been tested so far, but we can still improve our understanding of the drivers and constraints for changing clinicians’ behavior.

HOW CAN WE CHANGE BEHAVIOR?

Despite the availability of many methodologically sound studies our understanding of the processes involved in changing the behavior of healthcare professionals has not led to the definition of fail-proof interventions. A general consensus has emerged that there is no one ‘magic bullet’ for changing professional practice, and that interventions must be targeted at many levels to reach long lasting results, recognizing that the effects of improvement strategies will be context dependent. The variability in effectiveness of interventions, which sometimes work in one setting, but not in another one, is not fully understood. A comprehensive plan, targeting a range of problems and barriers to change, with strategies aimed at different levels - professional, team, patient, and organization- seems most likely to achieve long lasting changes in clinical behaviour.

The existing evidence suggests that interventions targeted at specific obstacles to change are more effective than generic interventions. Bosch and colleagues concluded that there is often a mismatch between the level of identified barriers and the type of interventions selected for an improvement strategy. Better tailoring of strategies to the identified level of being a barrier, should deserve more attention.

Michie and colleagues developed a theoretical, consensus-based framework that could
be used to select obstacles when designing an implementation strategy. They identified twelve domains to explain behavior change: knowledge, skills, social or professional role and identity, beliefs about capabilities, beliefs about consequences, motivation and goals, memory, attention and decision processes, environmental context and resources, social influences, emotion regulation, behavioral regulation and nature of the behavior.\(^9\)

Multiple concurrent interventions targeted on different domains should be selected, preferably with a specific goal of overcoming one or more of the barriers.\(^{11}\) However, methods used to identify impediments and tailor interventions to address them need further development.\(^{12}\)

**HOW SHOULD WE INTEGRATE RECRUITMENT OF PATIENTS INTO CLINICAL PRACTICE**

It is difficult not to agree on the large potential that exists for improving recruitment for trials at clinical sites.\(^4\) Unfortunately, we failed to identify a consistent pattern of trial characteristics associated with successful recruitment, though this failure is in line with previous, comparable attempts published in literature.\(^2,5\) This suggests that we should discontinue our search for simple, generally applicable predictors or ‘magic bullets’ that can guarantee success or failure in enrolling a sufficient number of patients.

A source of inspiration for alternative approaches could be the strategies that have been developed for improving implementation of evidence-based health care interventions in clinical practice. Several such strategies work with a series of plan-do-check-act cycles, in which one monitors whether a thoroughly designed recruitment plan, tailored to the trial, recruiting institute, patient population and expected obstacles, achieves the defined target. If not, one adjusts it further, as needed.\(^{13}\) McDonald and colleagues, propose the use of a business model approach and marketing techniques for recruitment in clinical trials. The model seems promising, but more examples of its application in practice are needed.\(^{14}\)

Expected obstacles could be derived from interviews and focus groups with patients and health professionals, ran before the study is developed, and from pilot studies. A review by Fletcher and colleagues on improving recruitment activity of clinicians in randomized controlled trials identified the use of qualitative methods to identify and overcome barriers to clinician recruitment activity as the most promising intervention.\(^{15}\) Further improving our knowledge about the processes of trial recruitment could further facilitate identification and elements that should be taken into account. Additionally, we should build strategies to convince and remind both health professionals and the general public of the value of research for practice. Familiarity and awareness of clinical research in the
general population could be enhanced by national public campaigns to improve awareness of clinical research.\textsuperscript{16,17}

The problems in recruiting participants also invokes a discussion on whether or not recruitment should be a shared societal responsibility, with obvious and unavoidable consequences for all clinicians and healthcare organizations. In our view, running and participating in trials should no longer be viewed as an optional activity, in one can opt in, but as an inescapable consequence of being a health professional and an absolute necessity for delivering quality care. Running trials and integrating trial results into clinical practice would then also be a responsibility for professional organizations, nationally and internationally. The rewards for participation in trials would then also change, from scientific recognition through authorship of articles, the now dominant form, to an essential element in the salary of individuals and the budget for healthcare organizations. The current opt-in system of inviting patients to trials could be replaced by an opt-out system for low risk evaluation research, based on the societal benefits of such a system. Such a change has already been proposed and tested in 2005 by Junghans and colleagues, where an opt-in system resulted in lower response rates and a biased sample.\textsuperscript{18} An opt-out system would not only improve participation rates, but could also lower the burden of patients, by shifting some of the responsibility of the often difficult decision making process about trial participation from patient to physician.\textsuperscript{19} Bernabe and colleagues argued that an opt-out procedure in minimal risk phase IV studies could be ethically acceptable.\textsuperscript{20} Crombie added to this discussion that “research should be undertaken only when there is a high likelihood of producing valid findings. Ethics requirements which result in invalid research may themselves be unethical”.\textsuperscript{21} The use and further investigation of such a system is also recommended by Al-Shahi Salman, in order to increase the value in biomedical research.\textsuperscript{22}

Patients could be informed about this general policy as soon as they enter the hospital, and are invited to sign a general informed consent about the use of data and efforts to improve quality. Such a system, has been pilot tested under federal regulations for clinical trials in pediatric medicine that are judged by the Institutional Review Board to reduce or have no effect on patient risk.\textsuperscript{23}

\textbf{HOW SHOULD WE INTEGRATE EBM TEACHING INTO PRACTICE?}

Based on a review of the literature, Coomarasamy and Khan concluded that EBM teaching
should be moved from classrooms to clinical practice, if one wants to see genuine improvement in relevant patient outcomes.\textsuperscript{24} Young and colleagues argued, based on a meta-review, that EBM teaching and learning strategies should not only be clinically integrated, but also multifaceted and including assessment.\textsuperscript{25}

In these processes, evaluation and assessment remain quintessential. Different methods for assessment of either teaching or non-teaching interventions have been developed, focusing on knowledge, skills or actual EBM behavior.\textsuperscript{26,27} Although valid and reliable instruments are available to measure the effect of teaching on knowledge and skills, the ultimate goal of EBM teaching is to improve patient outcomes, by changing behavior of the health professionals.

Our review on original studies that described the development or use of tools to assess EBP behavior in practice resulted in a variety of methods to assess EBM behavior: questionnaires, interviews or focus groups, observational studies or chart evaluations.\textsuperscript{27} Before one can select the optimal method for assessment, one has to be explicit about the objectives of such an assessment, and on what EBM should mean in practice.\textsuperscript{28} In our survey, TTT EBM course organizers highlighted the need for help in defining the curricula. How and with whom to build EBM curricula was also identified as an issue in a review for policy recommendations on how to improve EBM.\textsuperscript{29} A possible explanation for this might be that (clinical) EBM teachers are scattered over various (hospital) departments.

To our knowledge, there is at present no standard in medicine or in medical education about what we may expect from healthcare professionals when they are working evidence-based, and when we would qualify a health professional as working evidence-based in an adequate fashion. A survey amongst European EBM course organizers showed that most organizations welcomed a standardized European qualification in EBM.\textsuperscript{30} The Sicily statement (2005) postulates that “all health professionals need to understand the principles of Evidence-based Practice (EBP), recognize EBP in action, implement evidence-based policies, and have a critical attitude to their own practice and to evidence. Without these skills, professionals and organizations will find it difficult to provide “best practice”.\textsuperscript{31} Despite this broad definition, EBM training and assessment focuses often on searching and critical appraisal of original studies.\textsuperscript{27,30} We feel one can question whether these goals are the most relevant and attainable ones for trainees. Glasziou for instance showed, based on work from Kibbon et al and Hersh et al, that answers from students and doctors improved after a searching task, compared to before searching, but in 7 to 14 percent of cases answers went from right to wrong. In another 36 to 48 percent of cases initially wrong
still wrong, after searching. As training of healthcare professionals largely takes place in clinical practice, the question is who should teach EBM to trainees and, more importantly, who should be the role model for adequate EBM behavior in practice. Choudry et al. reported that 32 of 62 identified studies quality of care was lower for physicians with more years of experience, compared to more junior colleagues. This implies that more experienced clinicians may not necessarily be the best EBM role models. Constraints in resources and the complexity and vastness of the available evidence make it extremely unlikely that all healthcare professionals will become EBM experts, operating at a level in which they can identify and analyse the available evidence, evaluate validity and applicability of original articles, and apply the evidence appropriately in the care for individual patients. There is room for diversification in skills, in a structure where all health professionals understand and encourage EBM, have basic knowledge and skills (followers), while a selected group of health professionals, trained as EBM experts, gather, synthesize and apply the available evidence into evidence-based practice recommendations. Ubbink and colleagues proposed such a policy, where EBM should be exerted at micro level, meso level and macro level, supported by professional, educational and managerial role models.

Yousefi-Nooraie and colleagues discussed these issues with 51 EBP teachers from 15 countries and arrived at a consensus conclusion, that formulating clinical questions, searching pre-appraised resources, introduction to systematic reviews and critical appraisal of studies about therapy should be covered in starter level EBP courses, while other critical appraisal topics and quantitative decision-making techniques should be left to more advanced levels. If clinicians rely on - and depend on - pre-appraised resources, such as guidelines, valid and reliable pre-appraised sources and guidelines are even more necessary. Shaneyfelt pointed out that, in general, an improvement of the quality of guidelines is needed. A qualification system for clinical EBM experts, with requirements or expectations of the level of EBM knowledge and skills, could possibly help to reach and maintain high quality pre-appraised resources.

After more than 20 years, EBM has reached the age of maturity, but to make it fully functional, it has to be embedded in a coherent structure for teaching, managing and rewarding health care professionals. Only then will it deliver the anticipated gains in healthcare quality and efficiency, to the benefit of present and future patients.
REFERENCES

5. Oude Rengerink K, Bossuyt PM, Mol BW et al. Timely recruitment of the targeted number of participants perceived as more difficult than anticipated by many trialists – an analysis of a cohort of trials in the Netherlands Trial Register. *Unpublished*.
16. Reynolds T. Clinical trials: can technology solve the problem of low recruitment? Getting people to take part in clinical trials is often difficult. *BMJ* 2011;342.
32. Glasziou. CEBM. http://www.google.nl/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=1&ved=0CDUQFjAA&url=http%3A%2F%2Fwww.cebm.net%2Fmod_product%2Fdesign%2Fpresentations%2Fintroduction-to-ebm-2010-03.ppt&ei=8iXxUqfNuXw0gXB0IGIBQ&usg=AFQjCNEhQ05hTgfbMYq8-__4tDVACbbPt2RQ&sig2=HypK-BMrQNukyiDxvnGvHg&bvm=bv.60444564,d.d2k
CHAPTER 11

SAMENVATTING EN ALGEMENE BESCHOUWING
Dit proefschrift beschrijft een aantal onderzoeksprojecten rondom ‘evidence-based medicine’ (EBM), ofwel ‘op bewijslast gebaseerde geneeskunde’. In the projecten hier beschreven, richten we op zowel het genereren van bewijs, door klinische trials, als het integreren van gedegen bewijs in de klinische praktijk.

Dit proefschrift bestaat uit twee onderdelen. Deel 1 richt zich op het verbeteren van het rekruteren van het noodzakelijke aantal patiënten in klinische trials, aangezien dit een groot probleem is bij het evalueren van de effectiviteit van interventies in de zorg. Om ons begrip van het rekruteren van patiënten te verbeteren, probeerden we belemmerende en bevorderende factoren voor succesvol rekruteren te identificeren. De studies beschreven in deel 2 van dit proefschrift gaan in op het verbeteren van het integreren van op bewijslast gebaseerde besluitvorming in de klinische praktijk. We identificeerden barrières voor het onderwijzen van EBM in de klinische praktijk, ontwikkelden een EBM cursus ‘Train de trainer’ (TTT), en evalueerden methoden om EBM in de klinische praktijk te meten.

In Hoofdstuk 1 wordt de motivatie voor dit proefschrift beschreven. ‘Evidence-based Medicine’ (EBM), gedefinieerd in 1992 door Guyatt als ‘het integreren van klinische expertise met het best beschikbare bewijs uit systematisch onderzoek’, wordt gezien als de bouwsteen voor het maken van beslissingen in de huidige geneeskunde. De manier waarop we intuïtief kennis afleiden uit ervaring kan leiden tot vertekende beslissingen, daarom hebben we valide bewijs nodig van goede kwaliteit klinische trials.¹

DEEL 1: VERBETEREN VAN REKRUTEREN VAN PATIENTEN IN KLINISCHE TRIALS

In hoofdstuk 2 wordt de motivering voor en de opzet van de IMproving PArticipation of patients in Clinical Trials studie (IMPACT studie) beschreven. Om belemmerende en bevorderende factoren voor het succesvol rekruteren in trials beter te begrijpen vanuit verschillende niveaus, planden we twee cohort studies (hoofdstuk 3 en hoofdstuk 4) en een patiëntcontrole studie (hoofdstuk 5).

Hoofdstuk 3 beschrijft factoren die het succesvol rekruteren in klinische trials beïnvloeden, gebaseerd op een analyse van een cohort van 1129 trials geregistreerd in het Nederlands Trial Register met een geplande datum van het complementeren van het vooraf gestelde aantal patiënten tussen 2005 en 2010. Van 386 trials werden vragenlijsten ingevuld, deze rapporteerden over in totaal 242.412 deelnemers in meerdere specialismen, zowel in industrie als niet industrie-gesponsorde trials. Twee derde (66%) van de trial coördinatoren
gaf aan dat rekruteren moeilijker was dan verwacht. In minder dan de helft (43%) van de trials werd 80% of meer van het beoogde aantal patiënten gerekruteerd op de geplande einddatum, hoewel 53% de beoogde steekproef behaalde op de daadwerkelijke stopdatum. Slechts 52% van de trials werd gerapporteerd in de literatuur ten minste twee jaar na de geplande stopdatum van rekruteren. Het lukte ons niet om een consistent patroon van karakteristieken te identificeren die geassocieerd zijn met langzame rekrutering.

*Hoofdstuk 4* beschrijft de resultaten van een online vragenlijst, gestuurd naar lokale onderzoekscoördinatoren in centra die in totaal 14.808 patiënten rekruteerden voor 17 multicenter trials uitgevoerd binnen het Nederlandse Consortium voor Verloskunde & Gynaecologie studies in Nederland. We bestudeerden de associatie tussen kenmerken in een centrum en het gerekruteerde aantal patienten. We keken naar: de beslissing van een centrum over deelname aan trials, de onderzoeksoriëntatie van de afdeling en de (ervaren) logistieke ondersteuning door onderzoeksmedewerkers. In univariabele analyse was deelnemen aan trials met als reden ‘verwachten dat anderen rekruteren voor hun studies’ geassocieerd met hogere rekruteringsscores, deelnemen ‘omdat dat van ons verwacht wordt’ was geassocieerd met lagere scores. Hogere rekruteringsscores werden gezien in centra die regelmatig onderzoek initiëren en in derdelijns zorg centra. De gewogen rekruteringsscores waren hoger in centra waar onderzoeksmedewerkers meer uren beschikbaar waren.

In *hoofdstuk 5* beschrijven we de resultaten van 21 semi-gestructureerde interviews met zwangeren die uitgenodigd zijn voor deelname aan een gerandomiseerde studie in de verloskunde. We bestudeerden twee groepen: vrouwen die deelgenomen hebben (n=12) en vrouwen die deelname weigerden (N=9). We interviewden patiënten van 8 verschillende trials. Deelname aan wetenschappelijk onderzoek was voor 5 van de 12 deelnemers de belangrijkste reden om deel te nemen aan de studie, terwijl 5 noemden te hebben meegedaan omdat de interventie niet beschikbaar was buiten de trial. Belangrijke redenen om niet deel te nemen (n=9) waren een negatieve associatie of afkeer van de interventie, omdat het schade zou kunnen aanrichten (n=6) of vanwege praktische redenen (n=3). We identificeerden zeven hoofdthema’s die de beslissing om deel te nemen beïnvloedden. We concludeerden dat onzekerheid over wetenschappelijk onderzoek en/of de interventie aanzienlijk van belang was. Het begrip van patiënten omtrent wetenschappelijk onderzoek was over het algemeen
laag. Een persoonlijk, goed getimed en complete dialoog zou een gebalanceerde beslissing omtrent trial deelname kunnen verbeteren.

_Hoofdstuk 6_ laat resultaten zien van een vragenlijstonderzoek gestuurd naar cliniërs werkzaam op alle afdelingen Verloskunde en Gynaecologie in Nederland. We observeerden dat het rekruteren voor multicenter trials geassocieerd was met betere kennis van de resultaten van de trial. Alle trials meegenomen was het bij cliniërs die rekruteerden voor trials 1.8 keer meer waarschijnlijk dat ze de trial resultaten kenden, vergeleken met cliniërs die niet hadden gerekruiteerd (95% betrouwbaarheidsinterval (BI) 1.7 tot 1.9); variërend tussen 1.1 en 3.3 voor de verschillende studies. Het rekruteren voor een trial was ook significant geassocieerd met implementatie van de studieresultaten, hoewel het effect klein was, met een gepoold relatief risico van 1.1 (95% BI 1.02-1.19).

**DEEL 2: VERBETEREN VAN EVIDENCE-BASED MEDICINE IN DE KLINISCHE PRAKTIJK**

Er worden tegenwoordig veel cursussen over EBM aangeboden, toch blijft het een uitdaging hoe het lesgeven in EBM het beste in de klinische praktijk geïntegreerd kan worden. Zogenaamde Train de Trainer Evidence-based Medicine cursussen (TTT EBM), gericht op het onderwijzen van docenten van EBM, zouden kunnen helpen om de kwaliteit van het lesgeven in EBM te verbeteren en een standaard te creëren voor het lesgeven in EBM.

In het project beschreven in _hoofdstuk 7_ evalueerden we de beschikbaarheid en inhoud van TTT EBM cursussen. Via land specifieke EBM netwerken en infrastructuur identificeerden we 16 cursussen die als TTT EBM cursus werden aangemerkt: 4 in Hongarije, 4 in Nederland, 4 in het Verenigd Koninkrijk en 4 in Duitsland. In 15 van de 16 cursussen werd lesgegeven door academische specialisten en/of in 11 door methodologen of statistici. Zeven van de cursus organisatoren noemden de behoefte aan hulp met praktische voorbeelden van succesvolle technieken voor het lesgeven van EBM in de klinische praktijk, 7 behoefte aan hulp met het bepalen van het curriculum en 7 voor hulp met de financiering van TTT EBM cursussen.

_Hoofdstuk 8_ laat de resultaten zien van een evaluatie van een e-learning cursus voor klinische trainers om zelfverzekerd de EBM principes in de dagelijkse praktijk te onderwijzen.
De e-learning cursus bestond uit modules die de leerdoelen definiërenden en waarin met korte videoclips praktische en effectieve EBM onderwijsmethoden werden gedemonstreerd, voor 6 verschillende klinische settings. In een groep van 56 klinische tutoren in verschillende specialismen in zes Europese landen, verbeterden 47 (84%) hun score matig op een vragenlijst die ontworpen was om kennis te meten. De gemiddelde score voor de cursus was 69 (SD=10), deze nam toe tot 77 (SD=12) na de cursus (p<0.0001).

Maar, kennis over hoe les te geven in EBM leidt niet automatisch tot meer of beter lesgeven in EBM in de klinische praktijk. In hoofdstuk 9 identificeerden we en vergeleken we belemmerende en motiverende factoren voor het lesgeven in EBM in de klinische praktijk in verschillende Europese landen, op basis van een vragenlijst ingevuld door 120 klinische EBM onderwijzers uit 11 landen. Op een 7 punt Likert schaal, variërend van “1: helemaal geen belemmerende factor” aan het ene uiteinde tot “ 7: een onoverkomelijke belemmerende factor” aan de andere kant, kreeg “gebrek aan tijd” de hoogste score (mediaan 5). Matige belemmerende factoren (mediaan=4) waren “gebrek aan vereisten voor EBM vaardigheden” en “een piramidenvormige gezondheidszorg management structuur”. In Duitsland, Hongarije en Polen waren het lezen en begrijpen van artikelen in het Engels een hogere barrière dan in de andere onderzochte landen.

Zelfgerapporteerde obstakels en stimulerende factoren kunnen helpen om strategieën te bepalen die deze belemmerende factoren verminderen of bevorderende factoren versterken. Om te kunnen meten of zulke strategieën daadwerkelijk impact hebben op het gedrag in de praktijk, is een valide en betrouwbaar meetinstrument nodig. In hoofdstuk 10 presenteren we een systematisch literatuur overzicht van originele studies die de ontwikkeling of het gebruik van instrumenten die het evidence-based practice (EBP) gedrag onder alle gezondheidszorg professionals meten.

We identificeerden 172 studies: 117 vragenlijsten, 10 interviews of focus groepen, 9 observationele studies, 27 evaluaties van medische statussen en 9 studies die meerdere methoden combineerden. De psychometrische eigenschappen van de vragenlijsten die gebruikt waren werden gerapporteerd in ongeveer de helft van de studies. In 7 studies die een enkele EBM stap beoordeelden en in 6 studies die een combinatie van EBM stappen beoordeelden, werden zowel validiteit als betrouwbaarheid beschreven. Eén van deze vragenlijsten beoordeelde alle 5 stappen van EBM. We concludeerden dat valide en
betrouwbare methoden voor het meten van EBM zijn beschikbaar zijn, maar de keuze van de optimale methode voor het beoordelen van EBM gedrag in de praktijk, zou geleid moeten worden door het specifieke doel van de evaluatie.

IMPLICATIES VOOR DE PRAKTIJK EN SUGGESTIES VOOR VERDER ONDERZOEK
De studies beschreven in dit proefschrift en vergelijkbare projecten door anderen hebben laten zien dat het rekruteren van voldoende patiënten en het verbeteren van het met bewijslast onderbouwde klinische beslissingen maken complexe, context afhankelijke processen zijn.  

Over het algemeen zagen we dat het rekruteren van patiënten en EBM beter floreerden wanneer gezaaid in een vruchtbare grond: een milieu waar het rekruteren voor trials en het evidence-based werken (cultureel) ingebed zijn in de klinische praktijk en geaccepteerd worden, of zelfs verwacht, door opinieleiders en klinische trainers.  

Een vruchtbare grond is op zichzelf niet een voldoende voorwaarde voor clinici om te veranderen van een meer op autoriteit gebaseerde cultuur naar een cultuur waar de motivatie voor beslissingen kritisch beoordeeld wordt en open bediscussieerd.

Het uitnodigen van patiënten om mee te doen aan een trial verondersteld de moed om toe te geven dat de medische gemeenschap op dit moment niet weet wat het beste is, terwijl EBM direct of indirect de traditionele autoriteit van trainers en clinici uitdagaat.  

Een groot aantal implementatie strategieën is nu getest, maar we kunnen nog steeds ons begrip vergroten door te kijken naar stimuli en remmende factoren voor het veranderen van gedrag van klinici.

HOE KUNNEN WE GEDRAG VERANDEREN?
Ondanks de beschikbaarheid van veel methodologisch gedegen studies, heeft ons begrip van het proces rondom het veranderen van het gedrag van gezondheidszorg professionals niet geleid tot het opstellen van onfeilbare interventies. Er is algemene consensus ontstaan dat er geen eenvoudige oplossing is voor het verbeteren van de professionele praktijk, en dat interventies gericht zouden moeten worden op verschillende niveaus om lang aanhoudende effecten te bewerkstelligen, gegeven dat het effect van verbeter strategieën context afhankelijk is. De variabiliteit in de effectiviteit van interventies, die soms werken in één setting, maar niet in een andere, wordt niet volledig begrepen. Een verbeterplan dat verschillende problemen en barrières voor verandering aanpakt, met
strategieën gericht op verschillende levels – professional, team, patiënt en organisatie – lijkt het meest waarschijnlijk te leiden tot langdurige veranderingen in klinisch gedrag.\textsuperscript{6,10}

Het bestaande bewijs suggereert dat interventies gericht op specifieke obstakels voor verandering effectiever zijn dan algemene interventies.\textsuperscript{6} Bosch en collega’s concludeerden dat er vaak een mismatch is tussen de mate van de geïdentificeerde barrières en het type interventie dat geselecteerd wordt voor een verbeterstrategie.\textsuperscript{11} Het beter aanpassen van strategieën aan de geïdentificeerde mate van barrières, verdient daarom meer aandacht. Michie en collega’s ontwikkelden een theoretisch, op consensus gebaseerd raamwerk dat gebruikt kan worden voor het selecteren van obstakels bij het ontwikkelen van een implementatie strategie. Zij identificeerden 12 domeinen die gedragsverandering kunnen verklaren: kennis, vaardigheden, sociale of professionele rol en identiteit, overtuiging van mogelijkheden, overtuiging over consequenties, motivatie en doelen, geheugen, aandacht en besluitvormingsproces, sociale invloeden, emotie regulatie, gedragsregulatie en de aard van het gedrag.\textsuperscript{8} Meerdere, gelijktijdige interventies gericht op verschillende domeinen zouden geselecteerd moeten worden, het liefst met een specifiek doel om één of meer van deze barrières weg te halen. Maar, methoden die gebruikt worden om belemmeringen te identificeren en interventies zo aan te passen dat deze adequaat aangepakt worden hebben verdere ontwikkeling nodig.\textsuperscript{12}

\textbf{HOE MOETEN WE HET REKRUTEREN VAN PATIENTEN IN DE KLINISCHE PRAKTIJK INTEGREREN}

Het is moeilijk om het niet eens te zijn met de potentie die bestaat voor het verbeteren van het rekruteren voor trials in klinische centra.\textsuperscript{4} Helaas is het ons niet gelukt om een consistent patroon van trialkarakteristieken te identificeren die geassocieerd zijn met succesvol rekruteren, hoewel dit in lijn is met eerdere, vergelijkbare pogingen gepubliceerd in de literatuur.\textsuperscript{2,5} Dit suggereert dat we waarschijnlijk moeten ophouden met het zoeken naar eenvoudige, algemeen toepasbare voorspellers of oplossingen die het succes of falen van het rekruteren van een voldoende aantal patiënten kunnen garanderen.

Een bron van inspiratie voor alternatieve benaderingen zouden strategieën kunnen zijn die ontwikkeld zijn voor het verbeteren van de implementatie van op bewijskracht gestoelde gezondheidszorginterventies. Verschillende van deze strategieën werken met een serie plan-do-check-act cycli, waarin men monitort of een zorgvuldig ontwikkeld rekruteringsplan, aangepast aan de trial, het rekruterende centrum, de patiënten
Verwachte obstakels kunnen afgeleid worden uit interviews en focus groepen met patiënten en gezondheidszorg professionals, voordat de studie is gestart, en uit pilot studies. Een review door Fletcher en collega's naar het verbeteren van het rekruteren in gerandomiseerde, gecontroleerde trials identificeerde het gebruik van kwalitatieve methoden om barrières te identificeren en te verminderen als de meest veelbelovende interventie. Het verder verbeteren van onze kennis over het proces van het rekruteren in trials zou kunnen helpen bij het opsporen van andere factoren die van belang zijn. Aanvullend, zouden we strategieën moeten bouwen om zowel zorgprofessionals als de algemene populatie te overtuigen en herinneren aan de waarde van onderzoek voor de praktijk. Vertrouwdeheid met en bewustzijn van klinisch onderzoek in de algemene populatie zou verhoogd kunnen worden door nationale publieke campagnes om het besef van klinisch onderzoek te vergroten.

De problemen bij het rekruteren van deelnemers zet ook aan tot een discussie over of het rekruteren een gedeelde maatschappelijke verantwoordelijkheid is, met duidelijke en onvermijdbare consequenties voor alle clinici en gezondheidszorg organisaties. Vanuit ons oogpunt zou het uitvoeren en deelnemen in trials niet langer gezien moeten worden als een optionele activiteit, waarin iemand actief kan zeggen deel te nemen (opt-in), maar als een onontkoombare consequentie van het zijn van een zorgprofessional en een absolute noodzaak voor het leveren van kwalitatief hoogwaardige zorg. Het uitvoeren van trials en het integreren van trial resultaten in de klinische praktijk zou dan ook de verantwoordelijkheid moeten zijn van professionele organisaties, nationaal en internationaal. De beloning voor deelname aan trials zou dan ook veranderen, en een essentieel element worden in het salaris van de individuen en het budget van gezondheidszorg organisaties.

Het huidige opt-in systeem (waarbij een patiënt actief moet kiezen om deel te nemen) bij het uitnodigen van patiënten om deel te nemen aan trials zou vervangen kunnen worden door een opt-out systeem (waarbij een patiënt actief moet aangeven wanneer hij niet...
mee wil doen) voor laag-risico evaluatie onderzoek, gebaseerd op de maatschappelijke voordelen van zo’n systeem. Een dergelijke verandering werd ook voorgesteld en getest in 2005 door Junghans en collega’s, waar een opt-in systeem resulteerde in minder participatie en een geselecteerde steekproef. Een opt-out systeem zou niet alleen de deelname verbeteren, maar kan ook de last voor de patiënten verminderen, bij het verplaatsen van een deel van de verantwoordelijkheid van het vaak lastige besluitvormingsproces over deelname aan trials van de patiënt naar de arts. Bernabe en collega’s redeneerden dat een opt-out procedure in fase 4 studies met een laag risico ethisch acceptabel kan zijn. Bernabe voegde aan deze discussie toe dat “onderzoek alleen gedaan zou moeten worden als het erg waarschijnlijk is dat er valide resultaten worden geproduceerd. Ethische vereisten die resulteren in niet valide onderzoek kunnen op zichzelf onethisch zijn”. Het gebruik en verder onderzoek van zo’n systeem wordt ook aanbevolen door Al-Shahi Salman, met als doel het verhogen van de waarde van biomedisch onderzoek.

Patiënten zouden geïnformeerd kunnen worden over dit algemene beleid, zodra ze het ziekenhuis binnenkomen, en dan uitgenodigd worden om een algemeen informed consent formulier over het gebruik van data en de inspanningen om de kwaliteit te verbeteren. Een dergelijk systeem is getest, onder de federale regelgeving, voor klinische trials in de kindergeneeskunde die door een medisch ethische commissie of beoordelingscommissie van een instituut beoordeeld werden als geen effect hebbend op het risico van patiënten.

HOE ZOUDEN WE HET LESGEVEN IN EBM IN DE PRAKTIJK MOETEN INTEGREREN?

Gebaseerd op een systematisch literatuuroverzicht, concludeerden Coomarasamy en Khan dat het lesgeven in EBM zou moeten verplaatsen van klaslokalen naar de klinische praktijk, als men een daadwerkelijke verbetering wil zien in patiënt uitkomsten. Young en collega’s beargumenteerden dat EBM onderwijs en leerstrategieën niet alleen geïntegreerd zouden moeten worden in de kliniek, maar ook uit meerdere componenten zou moeten bestaan, inclusief assessment. In deze processen blijven evaluatie en beoordeling essentieel. Verschillende methoden voor het beoordelen van onderwijs of niet-onderwijs interventies zijn ontwikkeld, gericht op kennis, vaardigheden of daadwerkelijk EBM gedrag. Hoewel valide en betrouwbare instrumenten beschikbaar zijn voor het meten van het effect van het lesgeven op kennis en vaardigheden, is het uiterste doel van het lesgeven in EBM het verbeteren van uitkomsten van de patiënt, door het veranderen van het gedrag van zorgprofessionals. Ons systematisch literatuuroverzicht van originele studies die de ontwikkeling van
verschillende instrumenten om EBP gedrag in de praktijk te meten, resulteerde in een
verscheidenheid aan methoden om EBM gedrag te bepalen: vragenlijsten, interviews of
focus groepen, observationele studies of evaluaties van medische statussen. Voor men
de optimale methode van beoordeling kan selecteren, moet het doel van een dergelijke
beoordeling duidelijk zijn, en ook wat EBM zou moeten betekenen in de praktijk. In
ons dwarsdoorsnede onderzoek, noemden TTT EBM cursus organisatoren de noodzaak
voor hulp in het vaststellen van curricula. Hoe en met wie EBM curricula gebouwd zouden
moeten worden, werd ook geïdentificeerd als een probleem in een literatuuroverzicht met
beleidsaanbevelingen over hoe EBM te verbeteren. Een mogelijke verklaring hiervan
can zijn dat (klinische) EBM onderwijzers verdeeld zijn over verschillende (ziekenhuis)
afdelingen.

Volgens onze kennis, is er tegenwoordig geen standaard in de geneeskunde of in het
medisch onderwijs over wat we mogen verwachten van gezondheidszorg professionals
wanneer ze evidence-based werken, en wanneer we een gezondheidszorgprofessional
zouden kwalificeren als adequaat evidence-based werkend. Een dwarsdoorsnede
onderzoek onder Europepe EBM cursus organisatoren liet zien dat de meeste
organisatoren een gestandaardiseerde Europese kwalificatie in EBM verwelkomen. Het
Sicilië standpunt (2005) formuleert dat “alle zorg professionals de principes van evidence-
based practice (EBP) zouden moeten begrijpen, EBP in actie herkennen, op bewijskracht
gebaseerd beleid zouden moeten implementeren, en een kritische houding hebben naar
hun eigen praktijk en naar bewijs. Zonder deze vaardigheden zullen professionals en
organisaties moeite hebben om “best practice” te leveren. Ondanks deze brede definitie,
focust EBM onderwijs en beoordeling vaak op het zoeken en kritisch beoordelen van
oorspronkelijke studies. We hebben het gevoel dat we ons zouden kunnen afvragen
of dit de meest relevante en de te bereiken doelen zijn voor studenten. Glasziou liet
bijvoorbeeld zien, gebaseerd op werk van Kibbon en collega’s en Hersh en collega’s, dat
antwoorden van studenten en dokters verbeterden na een zoek taak, vergeleken met voor
het zoeken. Maar in 7 tot 14 procent van de gevallen gingen de antwoorden van goed naar
fout. In nog eens 36 tot 48 procent van de gevallen bleven initiële fouten antwoorden na
het zoeken fout.

Aangezien de training van zorg professionals vooral plaatsvindt in de kliniek, is het de vraag
wie EBM studenten zou moeten onderwijzen en, belangrijker, wie het rolmodel voor het
adequaat lesgeven van EBM gedrag in de praktijk zou moeten zijn. Choudry en collega’s
rapporteerden dat in 32 van 62 geïdentificeerde studies de kwaliteit van zorg lager was bij artsen met meer jaren ervaring, vergeleken met meer junior collega’s. Dit impliceert dat meer ervaren clinici niet noodzakelijkerwijs ook de beste EBM rolmodellen zijn. Beperkte middelen en de complexiteit en omvangrijkheid van het beschikbare bewijs maken het extreem onwaarschijnlijk dat alle zorgprofessionals EBM experts worden, die werken op een niveau waarop ze het beschikbare bewijs kunnen identificeren en analyseren, de validiteit en toepasbaarheid van originele artikelen kunnen beoordelen, en deze bewijskracht optimaal kunnen gebruiken in hun zorg voor individuele patiënten. Er is ruimte voor variatie van vaardigheden, in een structuur waar alle zorg professionals EBM begrijpen en aanmoedigen, basis kennis en vaardigheden hebben (volgers), terwijl een selecte groep zorgprofessionals, opgeleid als EBM experts, beschikbare bewijskracht verzamelen, aggregeren en toepassen in evidence-based aanbevelingen voor de kliniek. Ubbink en collega’s stelden een dergelijk beleid voor, waar EBM uitgevoerd zou moeten worden in micro, meso en macro niveau, ondersteund door professionele, onderwijskundige en management rolmodellen.

Yousefi-Nooraie en collega’s bediscussieerden deze onderwerpen met 51 EBP onderwijzers uit 15 landen en kwamen tot een consensus dat het formuleren van klinische vragen, het zoeken in vooraf beoordeelde bronnen, introductie in systematische literatuuroverzichten en het kritisch beoordelen van studies over therapie behandeld zouden moeten worden in EBP cursussen voor beginners, terwijl andere kritische beoordelingen en kwantitatieve besliskundige technieken overgelaten zouden moeten worden aan de EBM experts. Als clinici vertrouwen op - en afhankelijk zijn van - vooraf beoordeelde bronnen, zoals richtlijnen, zijn valide en betrouwbare vooraf beoordeelde bronnen en richtlijnen zelf nog meer noodzakelijk. Shaneyfelt wees erop dat in het algemeen een verbetering van de kwaliteit van richtlijnen nodig is. Een kwalificatie systeem voor klinisch EBM experts, met vereisten of verwachtingen van het level van EBM kennis en vaardigheden, kan mogelijk helpen tot het behalen en behouden van hoge kwaliteit bronnen met vooraf beoordeelde literatuur.

Na meer dan 20 jaar heeft EBM de volwassenheid bereikt, maar om het volledig functioneel te maken, zou het ingebed moeten worden als een coherente structuur voor onderwijs, management en beloning van zorg professionals. Alleen dan zal het de verwachte voordelen in kwaliteit van zorg en efficiëntie kunnen waarmaken, ten voordele van huidige en toekomstige patiënten.
REFERENCES

5. Oude Rengerink K, Bossuyt PM, Mol BW et al. Timely recruitment of the targeted number of participants perceived as more difficult than anticipated by many trialists – an analysis of a cohort of trials in the Netherlands Trial Register. Unpublished.
16. Reynolds T. Clinical trials: can technology solve the problem of low recruitment? Getting people to take part in clinical trials is often difficult. BMJ 2011;342.


32. Glasziou. CEBM. http://www.google.nl/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=1&ved=0CDUQFjAA&url=http%3A%2F%2Fwww.cebm.net%2Fdesign%2Fpresentations%2Fintroduction-to-ebm-2010-03.ppt&ei=8iXxUqfFfN0w0gXb0IGIBQ&usg=AFQjCNsEhQ05hTgbMYq8-4tDVACbbPt2RQ&sig2=HypK-BMrQNukyiDxvnGvHg&bvm=bv.60444564,d.d2k


37. Shaneyfelt T. In guidelines we cannot trust. Comment on ‘failure of clinical practice guidelines to meet institute of medicine standards’. Arch Intern Med 2012;172:1633-34
LIST OF CO-AUTHORS 
AND THEIR AFFILIATIONS
<table>
<thead>
<tr>
<th>Name</th>
<th>Affiliation</th>
</tr>
</thead>
</table>
| Chantal Arditi             | Institute of Social and Preventive Medicine  
                                Centre Hospitalier Universitaire Vaudois and University of Lausanne  
                                Switzerland                                                   |
| Theodoris Arvanitis        | School of Electronic, Electrical and Computer Engineering  
                                University of Birmingham, Birmingham  
                                United Kingdom                                                |
| Gemma Barnfield            | Department of Obstetrics and Gynaecology  
                                Birmingham Women’s NHS Foundation Trust  
                                United Kingdom                                               |
| Kitty W.M. Bloemenkamp     | Department of Obstetrics  
                                Leiden University Medical Center, Leiden  
                                The Netherlands                                              |
| Noortje van den Boogaard   | Department of Obstetrics and Gynaecology  
                                Academic Medical Center, Amsterdam  
                                The Netherlands                                              |
| Bernard Burnand            | Institute of Social and Preventive Medicine  
                                Centre Hospitalier Universitaire Vaudois and University of Lausanne  
                                Switzerland                                                   |
| Patrick M.M. Bossuyt       | Department of Clinical Epidemiology, Biostatistics and Bioinformatics  
                                Academic Medical Center, Amsterdam  
                                The Netherlands                                              |
| Nora A. Danhof             | Department of Obstetrics and Gynaecology  
                                Academic Medical Center, Amsterdam  
                                The Netherlands                                              |
| Nynke van Dijk             | Department of General Practice/ Family Medicine  
                                Academic Medical Center, Amsterdam  
                                The Netherlands                                              |
| Johannes J. Duvekot        | Department of Obstetrics and Gynaecology  
                                Erasmus University Medical Center, Rotterdam  
                                The Netherlands                                              |
| Jim van Eyck               | Department of Obstetrics and Gynaecology  
                                Isala Clinics, Zwolle  
                                The Netherlands                                              |
| Harry Gee                  | Birmingham Women’s Hospital  
                                Birmingham  
                                United Kingdom                                               |
| Birgit van der Goes        | Department of Obstetrics and Gynaecology  
                                Academic Medical Center, Amsterdam  
                                The Netherlands                                              |
| Monique C. Haak            | Department of Obstetrics and Gynaecology  
                                VU Medical Medical Center, Amsterdam  
                                The Netherlands                                              |
| Lotty Hooft                | Dutch Cochrane Center  
                                Academic Medical Center, Amsterdam  
                                The Netherlands                                              |
<table>
<thead>
<tr>
<th>Author</th>
<th>Affiliation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andrea R. Horvath</td>
<td>Department of Laboratory Medicine / TUDOR EBM Network</td>
</tr>
<tr>
<td></td>
<td>University of Szeged, Szeged</td>
</tr>
<tr>
<td></td>
<td>Hungary</td>
</tr>
<tr>
<td>Khalid S. Khan</td>
<td>Women’s Health Research Unit and Multidisciplinary Evidence Synthesis Hun (mEsh)</td>
</tr>
<tr>
<td></td>
<td>Queen Mary University of London, London</td>
</tr>
<tr>
<td></td>
<td>United Kingdom</td>
</tr>
<tr>
<td>Roy F.P.M. Kruijtwagen</td>
<td>Department of Obstetrics and Gynaecology</td>
</tr>
<tr>
<td></td>
<td>Maastricht University Medical Center, Maastricht</td>
</tr>
<tr>
<td></td>
<td>The Netherlands</td>
</tr>
<tr>
<td>Regina Kunz</td>
<td>Basel Institute for Clinical Epidemiology</td>
</tr>
<tr>
<td></td>
<td>Basel</td>
</tr>
<tr>
<td></td>
<td>Switzerland</td>
</tr>
<tr>
<td>Rogier Litjens</td>
<td>Department of Obstetrics and Gynaecology</td>
</tr>
<tr>
<td></td>
<td>Maastricht University Medical Center, Maastricht</td>
</tr>
<tr>
<td></td>
<td>The Netherlands</td>
</tr>
<tr>
<td>Sabine L.M. Logtenberg</td>
<td>Department of Obstetrics and Gynaecology</td>
</tr>
<tr>
<td></td>
<td>Onze Lieve Vrouwen Gasthuis, Amsterdam</td>
</tr>
<tr>
<td></td>
<td>The Netherlands</td>
</tr>
<tr>
<td>Berit Meyerrose</td>
<td>EBM and Guideline Department</td>
</tr>
<tr>
<td></td>
<td>Agency for Quality in Medicine, Berlin</td>
</tr>
<tr>
<td></td>
<td>Germany</td>
</tr>
<tr>
<td>Ben Willem J. Mol</td>
<td>Department of Obstetrics and Gynaecology</td>
</tr>
<tr>
<td></td>
<td>Academic Medical Center, Amsterdam</td>
</tr>
<tr>
<td></td>
<td>The Netherlands</td>
</tr>
<tr>
<td>Rita Onody</td>
<td>Department of Laboratory Medicine, Hemostasis Division</td>
</tr>
<tr>
<td></td>
<td>University of Szeged, Szeged</td>
</tr>
<tr>
<td></td>
<td>Hungary</td>
</tr>
<tr>
<td>Brent C. Opmeer</td>
<td>Clinical Research Unit</td>
</tr>
<tr>
<td></td>
<td>Academic Medical Center, Amsterdam</td>
</tr>
<tr>
<td></td>
<td>The Netherlands</td>
</tr>
<tr>
<td>Martijn A Oudijk</td>
<td>Department of Obstetrics and Gynaecology</td>
</tr>
<tr>
<td></td>
<td>University Medical Center Utrecht, Utrecht</td>
</tr>
<tr>
<td></td>
<td>The Netherlands</td>
</tr>
<tr>
<td>Maria G. van Pampus</td>
<td>Department of Obstetrics and Gynaecology</td>
</tr>
<tr>
<td></td>
<td>University Medical Center Groningen, Groningen</td>
</tr>
<tr>
<td></td>
<td>The Netherlands</td>
</tr>
<tr>
<td>Martina M. Porath</td>
<td>Department of Obstetrics and Gynaecology</td>
</tr>
<tr>
<td></td>
<td>Maxima Medical Center, Veldhoven</td>
</tr>
<tr>
<td></td>
<td>The Netherlands</td>
</tr>
<tr>
<td>Marko J. Sikkema</td>
<td>Department of Perinatology and Gynaecology</td>
</tr>
<tr>
<td></td>
<td>Hospital Group Twente, Almelo</td>
</tr>
<tr>
<td></td>
<td>The Netherlands</td>
</tr>
<tr>
<td>Olanrewaju Sorinola</td>
<td>South Warwickshire NHS / University of Warwick, UK</td>
</tr>
<tr>
<td></td>
<td>Foundation Trust, Warwick</td>
</tr>
<tr>
<td></td>
<td>United Kingdom</td>
</tr>
<tr>
<td>Marc E. Spaanderman</td>
<td>Department of Obstetrics and Gynaecology</td>
</tr>
<tr>
<td></td>
<td>Radboud University Medical Center, Nijmegen</td>
</tr>
<tr>
<td>Name</td>
<td>Affiliation</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Katja Suter</td>
<td>Basel Institute for Clinical Epidemiology</td>
</tr>
<tr>
<td></td>
<td>Basel</td>
</tr>
<tr>
<td></td>
<td>Switzerland</td>
</tr>
<tr>
<td>Shakila Thangaratinam</td>
<td>Women’s Health Research Unit and Multidisciplinary Evidence Synthesis Hun (mESh)</td>
</tr>
<tr>
<td></td>
<td>Queen Mary University of London, London</td>
</tr>
<tr>
<td></td>
<td>United Kingdom</td>
</tr>
<tr>
<td>Dirk Ubbink</td>
<td>Department of Quality Assurance and Process Innovation</td>
</tr>
<tr>
<td></td>
<td>Academic Medical Center, Amsterdam</td>
</tr>
<tr>
<td></td>
<td>The Netherlands</td>
</tr>
<tr>
<td>Hester Vermeulen</td>
<td>Department of Quality Assurance and Process Innovation</td>
</tr>
<tr>
<td></td>
<td>Academic Medical Center, Amsterdam</td>
</tr>
<tr>
<td></td>
<td>The Netherlands</td>
</tr>
<tr>
<td>Jacek Walczak</td>
<td>CASPolska</td>
</tr>
<tr>
<td></td>
<td>Krakow</td>
</tr>
<tr>
<td></td>
<td>Poland</td>
</tr>
<tr>
<td>Tobias Weberschock</td>
<td>Institute for General Practice / Department of Dermatology</td>
</tr>
<tr>
<td></td>
<td>Hospital of the Goethe, Goethe-University, Frankfurt</td>
</tr>
<tr>
<td></td>
<td>Germany</td>
</tr>
<tr>
<td>Susanne Weinbrenner</td>
<td>EBM and Guideline Department</td>
</tr>
<tr>
<td></td>
<td>Agency for Quality in Medicine, Berlin</td>
</tr>
<tr>
<td></td>
<td>Germany</td>
</tr>
<tr>
<td>Anna Wilmenska</td>
<td>CASPolska</td>
</tr>
<tr>
<td></td>
<td>Krakow</td>
</tr>
<tr>
<td></td>
<td>Poland</td>
</tr>
<tr>
<td>Christine Willekes</td>
<td>Department of Obstetrics and Gynaecology</td>
</tr>
<tr>
<td></td>
<td>Maastricht University Medical Center, Maastricht</td>
</tr>
<tr>
<td></td>
<td>The Netherlands</td>
</tr>
<tr>
<td>Gianni Zanrei</td>
<td>Universita Cattolica del Sacro Curore</td>
</tr>
<tr>
<td></td>
<td>Piacenza</td>
</tr>
<tr>
<td></td>
<td>Italy</td>
</tr>
<tr>
<td>Sandra Zwolsman</td>
<td>Department of General Practice/ Family Medicine</td>
</tr>
<tr>
<td></td>
<td>Academic Medical Center, Amsterdam</td>
</tr>
<tr>
<td></td>
<td>The Netherlands</td>
</tr>
</tbody>
</table>
PORTFOLIO
## 1. PHD TRAINING

### A. COURSES (year; ECTS)

**Academic Medical Center, Amsterdam**
- Evidence based searching (2008; 0,1)
- AMC Basic world of science (2008; 0,7)
- Advanced biostatistics (2009; 2,1)
- Scientific writing in English (2009; 1,5)
- Advanced topics in clinical epidemiology (2009; 1,1)
- Systematic reviews (2009; 0,3)
- Advanced biostatistics (2010; 2,1)
- Qualitative health research (2010; 1,9)
- Clinical data management (2010; 0,3)
- Caste course clinical research (2010; 0,9)
- Genetic epidemiology (2010; 1,1)
- Basic course legislation and organisation for clinical researchers (2010; 0,9)
- Career development (2010; 0,8)
- Introduction in R (2010; 0,1)

**Leiden University Medical Center, Leiden**
- Mixed and longitudinal modeling (2010; 6,0)

**Erasmus University Medical Center, Erasmus Winter Program, Rotterdam**
- Advanced topics in clinical trials (2011; 1,9)

**Swiss Epidemiology Winter School, Institute of Social and Preventive Medicine, Bern**
- Indirect comparison and network meta-analysis (2014; 1,5)

**Julius Academy, University Medical Center Utrecht, Utrecht**
- Multivariate analysis (2012; 1,5)

### SEMINARS, WORKSHOPS AND MASTER CLASSES

- Aprove symposium 'Scientist get out of your lab' (2010), Amsterdam
- Workshop global maternal health (2011), San Francisco
- Workshop maternal medical complications (2011), San Francisco
- Master class GRADE by Gordon Gyatt (2011), Amsterdam
- Aprove symposium 'Train your brain' (2011), Amsterdam
- Workshop systematic review of instruments (2012), Amsterdam
- Workshop statistical issues in interim analysis (2012), Miami
- Workshop adaptive design in clinical trials (2012), Miami
- Workshop Data Safety Monitoring (2013), Utrecht
- Workshop sitting on or chairing a Data Safety Monitoring Board (2013), Boston
- Workshop using new media tools in clinical trials (2013), Boston
- Master class ‘Who wrote my paper’ by dr. Drummond Rennie (2013), Boston
- Weekly department lunch meetings (2008-2013), Amsterdam
- Weekly department seminars and journal club (2008-2013), Amsterdam
- Seminars Clinical Epidemiology, Biostatistics and Bioinformatics (2008-2013), Amsterdam
B. Presentations as first author

- Improving Participation of Clinical Trials, design and rationale of IMPACT. Werkgroep Epidemiologie Nederland (WEON), Nijmegen, June 2010. *Oral Presentation.*
- Pregnant women’s views about participation in trials – a qualitative study. Society for Clinical Trials (SCT), May 2012, Miami. *Poster presentation.*
- Incorporation of clinical trials in routine patient care – RCTs as the standard of care rather than the exception. Society for clinical trials, Boston. *Invited session. Oral presentation.*

C. (Inter)national conferences

- Werkgroep Epidemiologie Nederland (WEON), 2010, Nijmegen
- Society for Maternal Fetal Medicine (SMFM), 2011, San Francisco, US
- Society for Clinical Trials (SCT), 2012, Miami, US
- Werkgroep Epidemiologie Nederland (WEON), 2013, Utrecht
- Society for Clinical Trials (SCT), 2013, Boston, US.

D. Other

- Consortium Training Days, 2012, Veldhoven
- Castle Course Clinical Epidemiology, 2010, France
- EU EBM Unity partnership
- EBM Connect collaboration

2. TEACHING

Teaching
- Tutor Practical Biostatistics Course, Academic Medical Center, 2010-2011.

Supervising
- Sabine Logtenberg (2010). Topic: pregnant women’s motives for (non)participation in trials.
3. GRANTS

- Co-applicant of the ZonMw funded project on prevention of congenital cytomegalovirus infection, main applicant dr. E. Pajkrt.

4. LIST OF PUBLICATIONS


Bloemenkamp KWM, on behalf of the PROBAAT study group. **Cost-effectiveness of a Foley catheter compared with prostaglandin E2 gel for induction of labour at term (PROBAAT trial).** BJOG 2013;120:987-95.


Marta Jozwiak, Katrien Oude Rengerink, Marjan Benthem, Erik van Beek, Marja Dijkstra, Irene de Graaf, Marloes van Huizen, Jan Willem de Leeuw, Martijn Oudijk, Mariëlle van Pampus, Dimitri Papatsonis, Denise Perquin, Martina Porath, Joris van der Post, Robert Rijnders, Marc Spaanderman, Christine Willekes, Ben Willem Mol, Kitty Bloemenkamp. **Induction of labor at term in women with an unfavourable cervix: a comparison of Foley catheter and Prostin (PROBAAT trial).** Lancet. 2011;378(9809):2095-103.


Oude Rengerink, Katrien; Thangaratinam, Shakila; Barnfield, Gemma Alexandra; Suter, Katja; Horvath, Andrea R; Walczak, Jacek; Kaleta, Anna; Weinbrenner, Susanne; Meyerrose, Berit; Arvanitis, Theodoris N.; Onody, Rita; Zanrei, Gianni; Kunz, Regina; Arditi, Chantal; Burnand, Bernard; Gee, Harry; Khan, Khalid; Mol, Ben WJ. How can we teach Evidence Based Medicine in practice? Barriers to on-the-job EBM teaching and learning. *Medical Teacher.* 2011;33(3):e125-30.


FOR THE EU EBM UNITY (GROUP AUTHORSHIP)


FOR THE EBM CONNECT COLLABORATION (GROUP AUTHORSHIP)


DANKWOORD
Dankwoord

Onderzoek doe je gelukkig niet alleen. Een aantal mensen wil ik in het bijzonder bedanken.

In de eerste plaats wil ik de mensen bedanken die op diverse manieren (belangenloos) hebben bijgedragen aan de diverse onderzoeken: zonder deelname geen onderzoek. Voor het IMPACT project wil ik graag alle trialcoördinatoren, gynaecologen, arts-assistenten en verloskundigen bedanken voor het invullen van de vragenlijst over rekruteren in trials. Alle patiënten die meegewerkt hebben aan een interview wil ik bedanken. Daarnaast wil ik ook alle onderzoekers binnen en buiten de verloskunde bedanken die deze studies mede mogelijk hebben gemaakt.

Ook wil ik alle gynaecologen bedanken die meegewerkt hebben aan het piloten van de e-learning modules, de EBM docenten uit diverse Europese landen die de vragenlijst over EBM barrières hebben ingevuld en alle respondenten op de inventarisatie van EBM TTT cursussen in Europa.

Mijn promotor Prof. dr. B.W.J. Mol. Ben Willem, ik wil je bedanken voor de kansen die je me geboden hebt en het vertrouwen dat je me hebt gegeven, vooral ook wat betreft toegang tot je uitgebreide (inter)nationale netwerk. Hoewel - of waarschijnlijk dankzij - in onze wekelijkse besprekingen niet altijd mijn artikelen centraal stonden, maar je alles vertelde over de perikelen in de Nederlandse gezondheidszorg of de wereldwijde perinatologische zorg, ga ik ze missen.

Mijn promotor Prof. dr. P.M.M. Bossuyt. Patrick, je werd mijn promotor toen mijn promotie al een tijdje gestart was, maar je bijdrage was niet minder belangrijk. Ik heb veel geleerd van je kritische blik. Veel dank voor je snelle hulp bij de afronding.


De leden van mijn promotiecommissie Prof. dr. R.J.P.M. Scholten, Prof. dr. F. van der Veen, Prof. dr. R.J. de Haan, Prof. dr. M.A. Boermeester, Prof. dr. G.P. Westert en Prof. dr. T.J. ten Cate wil ik bedanken voor het beoordelen van het manuscript en het plaatsnemen in de commissie.

Ik wil graag alle coauteurs die bijgedragen hebben aan een van de stukken in dit proefschrift bedanken.
I want to thank all co-authors of the EU EBM Unity, for collaboration in the Teach the Teacher EU EBM project. Working with so much experienced people from different fields of medicine and different countries was interesting and the meetings were enjoyable. Special thanks to Khalid Khan and Shakila Thangaratinam for initiating and leading the project.

Sandra, Nynke, Dirk en Hester, er leek geen eind te komen aan het review dat we samen hebben opgezet vanuit de gynaecologie, huisartsgeneeskunde en de KPI pijler EBM implementatie. Het was fijn om met jullie samen te werken. Nynke, dank voor het (last-minute) brainstormen over de discussie van het proefschrift.

Sabine, Liekje en Janna. Jullie hebben voor jullie stage een belangrijke bijdrage geleverd aan de interviews. Dank voor jullie enthousiasme!

Rogier, ik werd later bij het project over rekruteren en implementeren betrokken dat je al was gestart met Ben Willem. Dank voor onze samenwerking.

Norah, dank voor je bijdrage hieraan en aan het project over de METC procedure!

Ingrid, Marjan en Corien, dank voor jullie hulp. Van het starten als beginnende onderzoeker, het maken van afspraken, tot de laatste loodjes van dit proefschrift. Maya, dank voor al je hulp bij vragen over het consortium.

Noor, zowel voor jou als voor mijn proefschrift hebben we stukken samen geschreven. Door je combinatie van inzicht in onderzoek, kliniek en in mensen is het leuk om met je te brainstormen.

Brent, je was betrokken bij de opzet van het IMPACT project. Dank voor je input en discussies hierover. We spreken elkaar nog maandelijks bij het consortium methodologenoverleg.

Gynaecologen en arts-assistenten in het AMC en in het consortium. Van zo dichtbij meekijken hoe de klinische praktijk in elkaar zit heeft me veel geleerd. Dank daarvoor.

Collega onderzoekers in het AMC en in het Consortium, jullie zijn met teveel om op te noemen. Het is fijn om elke dag binnen zo’n enthousiaste club mensen te werken. Hoewel ik altijd een beetje vreemd wordt aangekeken, en inmiddels gewend ben aan de vraag:
Dankwoord

“wàt ga je dan doen als je geen gynaecoloog gaat worden?” De (vrijdagmiddag) borrels, bootje varen over de grachten, skien en de assistentenweekenden: altijd gezellig. Ik hoop dat diner Chique des Chercheurs een traditie gaat worden...
Parvin and Janneke, the Iranian dancing on the roof terrace in Boston was lovely.

Leukste Kamer app, Hannah, Sabra, Maaieke en Myrthe, gezellig om de dag met jullie (en een capu) te beginnen!

Emily, Jolande, Margreet en Floortje, lieve vissenkommers! De vissenkom bestond al toen ik me bij jullie voegde, maar dat maakte niet dat ik me minder welkom voelde. Lieve Jolande, hoewel ik soms echt gek werd van je geklets in de vissenkom, was het altijd gezellig om samen een biertje te drinken en over het onderzoek na te denken. Leuk dat je bent komen helpen schilderen in mijn nieuwe huis, kom snel weer eens langs! Lieve Emily, het was fijn om met jou op de kamer te zitten. De discussies over de aanpak van stukken waren leuk, en ik leerde er ook altijd van. Ik kom binnenkort je nieuwe huis bewonderen. Lieve Floor, we zijn samen twee keer paranimf geweest, dat was leuk. Als het goed is werk je nu aan de laatste loodjes, van je promotie en je zwangerschap. Lieve Margreet, het was altijd gezellig als we samen in de vissekom zaten, en relaxed om samen te werken. Leuk je nu als moeder te zien. Het liefst had ik jullie allemaal als paranimf, maar dat ging helaas niet. Snel weer eens eten!

Moniek, Inge en Annet. Het was gezellig regelmatig de week af te sluiten met een kabouterbiertje. Het Zuiderzee avontuur was memorabel, en niet alleen om de hoge golven. Moniek, er zijn niet veel mensen die ik graag om 7 uur tegenkom om op de fiets naar Breukelen te gaan, maar met jou is het altijd gezellig.

Er zijn daarnaast nog een heel aantal mensen die niet betrokken zijn geweest bij mijn onderzoek, maar wel belangrijk in mijn leven.

Lieve vrienden uit Nijmegen: Anneke, Sally, Thijs, Bark, Jeanine, Marcia, Noor, Laura, Ellen en Bas, en Renske. Lieve vrienden van het studentenorkest Sally, Ellen, Erica, Anneke, Suzan, Jacolien en Marit. Lieve vrienden uit Denekamp en omstreken Sandra, Femke, Ilon, Kim, Rian. Ik ben een beetje verder weg gaan wonen, maar gelukkig zie ik jullie nog regelmatig. Zonder jullie was het leven een stuk minder de moeite waard, bedankt voor alle gezelligheid!
Mijn paranimfen, Ellen en Janne. Lieve El, samen zijn we vanuit Nijmegen naar ons huisje aan de Balistraat vertrokken. Net als je vriendschap voelt het vanzelfsprekend dat je vandaag naast me staat. Wat we ook ondernemen samen, het gaat vaak zonder moeite en het wordt altijd leuk: St. Petersburg, Lapland, New York, skiën in Frankrijk of gewoon een kopje thee drinken. Ik ben trots dat je vandaag naast me staat!

Lieve Janne, sis, fijn dat je vandaag naast me staat. Iemand die zelf aanbied alvast een feest te organiseren (zodat de promotie dan wel af moet zijn...) en het boekje gaat ontwerpen moet je wel als paranimf vragen. We hebben samen twee mooie reizen gemaakt, Vietnam en Namibië. Fijn dat ik altijd bij je terecht kan.

Lieve Tiny, Wim, Marleen, Lucas en Marieke. Ik hoop nog veel gezellige dingen samen te doen.

Lieve oma, wat speciaal dat je dit op je 91e nog kunt meemaken. Sis en bro’s, lieve Janne & Daan, Jens & Carolien, Karel & Pien. Dank voor de vanzelfsprekendheid waarmee we er voor elkaar zijn. Lieve pap, wat fijn dat ik altijd op je kan rekenen, of het nou voor klussen is of voor iets anders. Binnenkort klussen bij jou. Lieve mam, aan het begin begreep je niet echt wat ik nu precies aan het doen was, maar er was altijd een luisterend oor en je begreep heel goed of ik het leuk vond of niet. Jullie lieten me altijd vrij om te doen wat mij goed leek en waren altijd bereid hierbij te helpen, bedankt. Mam, het wordt inderdaad tijd om weer wat meer muziek te maken.

Lieve René, ik ben gelukkig dat ik bij jou zo mezelf kan zijn en dat je zo oprecht van me houd. Wat fijn om met je samen te zijn!