Antenatal education – a systematic review and a randomised trial

Carina Sjöberg Brixval

National Institute of Public Health, Faculty of Health Sciences, University of Southern Denmark

PhD thesis submitted April, 2016.

Supervisors

Professor Pernille Due, DMSci, National Institute of Public Health, University of Southern Denmark.

Associate professor Lau Caspar Thygesen, PhD, National Institute of Public Health, University of Southern Denmark.

Senior researcher Vibeke Koushede, PhD, National Institute of Public Health, University of Southern Denmark.

Assessment committee

Professor in Obstetrics and Gynecology, MD, PhD Ulrik Schiøler Kesmodel, Institute of Clinical Medicine, University of Copenhagen and Herlev University Hospital, Denmark.

Associate professor, MD, PhD Anna Sarkadi, Department of Women's and Children's Health, Uppsala University, Sweden.

Associate professor, LM, MPH, PhD Dorte Hvidtjørn (chair of the committee), Institute of Clinical Research, University of Southern Denmark.

PREFACE AND ACKNOWLEDGEMENTS

The work presented in this thesis was carried out at the National Institute of Public Health, University of Southern Denmark. I joined the research team behind the NEWBORN trial as a research assistant in 2011 while the project was still in the planning phase. In the following four years, we planned the trial, developed the course material, recruited and randomised participants, collected questionnaire data, and finally analysed data. It has been a long, but very fruitful and fun process and I have learned a lot about intervention research during these years.

I have been collaborating with a lot of people during my time as a PhD student and I owe great thanks to everyone who have contributed to my scientific education. First and foremost I would like to thank my always extremely supportive supervisors: Lau Caspar Thygesen, Vibeke Koushede, and Pernille Due. Thank you for the time and energy you have spent on supervising me. Each of you has contributed to my work with your special field of expertise and I believe I have had the best possible team to guide me through my PhD period.

Also, I wish to thank my co-authors. A special thanks to Solveig Forberg Axelsen for being a great colleague and for many hours of hard work on the NEWBORN trial. Thanks to my collaborators at the department of Obstetrics and Gynecology at Hvidovre Hospital: Tom Weber and Christina Rørbye. I also thank Nanna Roed Johansen for spending lots of hours on carefully going through the medical birth records while we conducted the validation of the Obstetric Database.

Thanks to all my colleagues at the National Institute of Public Health for an inspiring research environment and a friendly atmosphere – and for good talks while visiting the coffee machine.

From September to December 2015 I visited the Copenhagen Trial Unit, Rigshospitalet, where I spent four months with fantastic colleagues. I wish to especially thank Christian Gluud for giving me the opportunity to join the CTU team, for great inspiration, and for sharing his enthusiasm for systematic reviews and clinical trials. Also thanks to all the other staff members at CTU.

Finally, thanks to my children, Alexander and Benedicte, and especially to Christian for always being supportive and for patiently listening to my scientific ups and downs.

Carina Sjöberg Brixval
Copenhagen, April 2016

TABLE OF CONTENTS

1. INTRODUCTION

In 1972 the British epidemiologist, Archie Cochrane, emphasised that as health care resources will always be limited, resources should be used to provide the forms of health care proven to be most effective. According to Cochrane the decision about what care to provide, should be based on results from the most reliable source of evidence, namely randomised trials (1).

Later in 1992 the term "evidence-based medicine" was introduced (2). In line with Cochrane's recommendations evidence-based medicine is intended to optimise decision-making by emphasising use of evidence, preferably from well designed and conducted research in the form of meta-analyses, systematic reviews, and randomised trials, to guide recommendations for practise. Decisions on provision of care should hence be based on firm evidence, rather than trends and personal beliefs of practitioners, experts, or administrators.

Systematic reviews attempt to gather all the empirical evidence that fits pre-specified criteria in order to answer a specific research question (3). The aim is to minimise bias by the use of systematic methods and provide more robust evidence for decision making than what can be gathered from singular randomised trials (3). A main advantage of systematic reviews is that they synthesise the research-based evidence and present the results in an accessible format, meaning that healthcare providers and policy makers are able to base decisions on provision of healthcare without having to review and assess vast amounts of information. The extent to which a review can be used to draw conclusions on the effect of a given intervention depends on the methodological quality of the review itself (4) as well as the validity of the included studies (3). The quality of systematic reviews can be heightened by following the recommendations from the Cochrane Collaboration (3).

Randomised trials are advantageous over observational studies when assessing effects of healthcare interventions (5). The idea behind the randomised trial is to make the groups comparable on all other parameters than the assigned treatment group, both in terms of known and unknown factors. This means that the effect estimate will be unbiased given that the trial is free of bias related to e.g. the randomisation process, blinding, and completeness of outcome data (3). It is therefore crucial that randomised trials of high quality are conducted and reported, and that they are included as evidence base of systematic reviews.

The overall aim of my PhD study was to assess the effect of antenatal education in small classes on use of pain relief, obstetric interventions, and childbirth self-efficacy based on evidence from undertaking a systematic review and analysing data from a randomised trial that I have contributed to develop, implement and evaluate; the NEWBORN trial.

2. BACKGROUND

2.1. ANTENATAL EDUCATION

Antenatal education aims to help prospective parents prepare for childbirth and parenthood. The education uses a range of educational and supportive measures to help parents understand social, emotional, psychological, and physical needs during pregnancy, labour and parenthood (6).

CHANGES OF ANTENATAL EDUCATION OVER TIME

In most Western countries antenatal education is well-established, but the form and content has changed markedly over time. As an example, practice has at certain points in time been centred on antenatal education in small classes with group discussions – at other points in time practice has been to offer expectant parents lectures in large auditoriums. These changes have occurred without evidence of an effect of specific types of antenatal education on relevant outcomes (6). Although antenatal education is widely used and hence represents considerable costs to the healthcare system, the effect of form and content of the education is poorly evaluated (6).

I Denmark, antenatal education was for many years provided in small classes. However, over a number of years from approximately 2005 and onwards several birth sites in the Capital Region changed their offer to large-scale auditorium-based lectures with a minimum of interaction with the audience. However, over the last two to three years, the Capital Regions' antenatal education offer has moved away from the auditorium-based lectures and back to antenatal birth and parent preparation classes in small groups for all expectant parents. The shifts in practice appear to be based on tendencies, professional beliefs, political wishes, and economic considerations rather than on solid evidence from trials favouring antenatal education in small classes over auditorium-based lectures.

In their recommendations on antenatal care from 2009, the Danish Health Authority declares, that little is known about what antenatal education should encompass in order to meet the needs of parents today (7). However, it is highlighted that several studies describe parents' wishes to discuss aspects related to the social, emotional, and psychological side of parenthood, and how to interact with their newborn baby, in addition to gaining information about the delivery. Therefore, they recommend that antenatal education comprises these aspects (7). These recommendations constituted the relevant background for the NEWBORN trial, which was developed in 2011. In 2013, new recommendations were issued (8). No changes were made to the section on recommendations for antenatal education between the two editions.

The Danish Health Authority additionally states that experience shows that group activities can contribute to the creation of network that the woman and her partner can benefit from after the delivery (7). Based on the advantages of the network creation and the increased possibility for the expecting parents to discuss

the birth and transition to parenthood in relation to practical, psychological, and social aspects in smaller groups, the Danish Health Authority recommends that antenatal education is delivered in small classes (9). However, antenatal education in small classes is more costly regarding expenses used for facilities and midwife salaries compared to auditorium-based lectures which may be an incentive for the hospitals to offer the auditorium-based approach. In 2011, none of the Danish regions provided birth and parent preparation in the form and extent recommended by the Danish Health Authority (9).

In 2011, when the NEWBORN trial was planned, antenatal education was still offered as auditorium-based lectures in the Capital Region of Denmark, although the Capital Region's Regional Birth Planning Committee at the time recommended implementation of antenatal education birth and parent preparation in small classes. The Capital Region of Denmark was therefore an ideal setting in which to conduct a randomised trial of antenatal education in small classes versus auditorium-based lectures.

ANTENATAL EDUCATION IN SMALL CLASSES

In recent time, principles of adult learning have been given more weight in antenatal education. It has been argued that information transfer by itself should no longer be the focus of antenatal education. Rather, all health-promotion should provide opportunities for people to learn skills in order to practice desired behaviours (10). Learning theorists suggest that educators need to become facilitators which shifts the emphasis from the educator to the learner (11). Additionally, it is emphasised that many individuals benefit from learning through an activity in which they become actively engaged. Further, it is highlighted that people learn more effectively in a group setting, where they have the opportunity to assume different roles, to observe others' perspectives, to interact regularly, and to supplement one another (12).

Qualitative research on pregnant women's preferences in relation to antenatal education has suggested that women want to be able to ask questions, seek clarification, and relate information to their own circumstances (13). They prefer antenatal education in small classes with participation of a small number of participants; and further, that the educator functions as a facilitator that promotes discussion, gives suggestions for practicing skills, and encourages participants to get to know and support one another (13).

Based on these considerations on adult learning principles and women's own preferences, antenatal education in small classes with participation of a limited number of expectant women and their partners, may create an environment which enables couples to discuss feelings and concerns. Further, education in small classes may enhance the participant's awareness of own resources and provide them with problem-solving strategies thereby promoting important competences to cope with pregnancy, birth and parenthood. These increased competences may lead to health care savings in the long term even though the expenses are larger for education in small classes than for auditorium-based education.

2.2. PAIN RELIEF AND OBSTETRIC INTERVENTIONS

Use of pain relief during labour as well as obstetric interventions are common. In 2013, epidural analgesia was used in 24% of all deliveries in Denmark, oxytocin administration was used in 22% of all deliveries; 7% of the infants were delivered by vacuum extraction, 12% by emergency caesarean section, and 10% by elective caesarean section (14).

Epidural analgesia provides effective pain relief, and compared to other types of pain relief, women using epidural analgesia report faster relief of pain and better pain relief in both the first and second stages of labour (15). Additionally, epidural analgesia has advantages over use of morphine. Babies of mothers who use epidural analgesia are at lower risk of requiring naloxone to block the effects of opioid compared to babies of mothers who use opioids, e.g. morphine, as pain relief (15). However, epidural analgesia has a range of potential side effects on the birth process. Longer second stage of labour and increased risk of interventions such as oxytocin administration, instrumental vaginal birth, e.g. vacuum extraction, and caesarean section are prevalent (15-22). Also, interventions in the birth process may have negative side effects on the infant. Use of epidural analgesia may lead to lower Apgar scores and hypotonia (20), vacuum extraction may result in facial and scalp injuries (23), and caesarean section increases the risk of respiratory morbidity in the infant (24). In addition, use of epidural analgesia and obstetric interventions are costly for the healthcare system (25).

ANXIETY, PAIN, AND EARLY ADMISSION

Anxiety during labour increase the level of the stress hormone epinephrine in the blood, which inhibits release of the contraction stimulating hormone; oxytocin. This may in turn lead to decreased uterine contractility and a longer active labour phase (26), which may result in exhaustion of the woman. Further, a woman's fear and anxiety influences her experience of pain (18), which may lead to increased use of pain relief, e.g. epidural analgesia (27).

Anxiety and insecurity may increase the likelihood that women request early admission to the labour ward although they are only in the latent phase of labour (28). Admission to hospital before having entered the period of active labour increases the risk of use of pain relief, e.g. epidural analgesia, obstetric interventions as well as complications (29-32). This mechanism is, among other things, due to the more frequent diagnosis of dystocia because of increased monitoring of labour and a "clinical cascade effect" once a woman is admitted to the hospital (33). It is therefore recommended that women wait until the active phase of labour before seeking admission, and that the health personnel encourage women to follow this recommendation (31).

2.3. CHILDBIRTH SELF-EFFICACY

According to Bandura, self-efficacy is concerned with judgments of how well one can execute courses of action required to deal with prospective situations (34). Two important and independent components of the theory of self-efficacy are; 1) outcome expectancy, i.e. the individual's belief that a given behaviour will lead to certain outcomes, and 2) efficacy expectations, i.e. the individual's belief in own ability to perform a specific behaviour. Further, Bandura operates with four main sources of self-efficacy, i.e. four different ways that a person's self-efficacy may increase: 1) personal experiences of mastering, 2) vicarious experiences provided by others, e.g. by hearing about or observing other people's experiences, 3) social persuasion by encouragement from others, and 4) emotional interpretations of physical states (35).

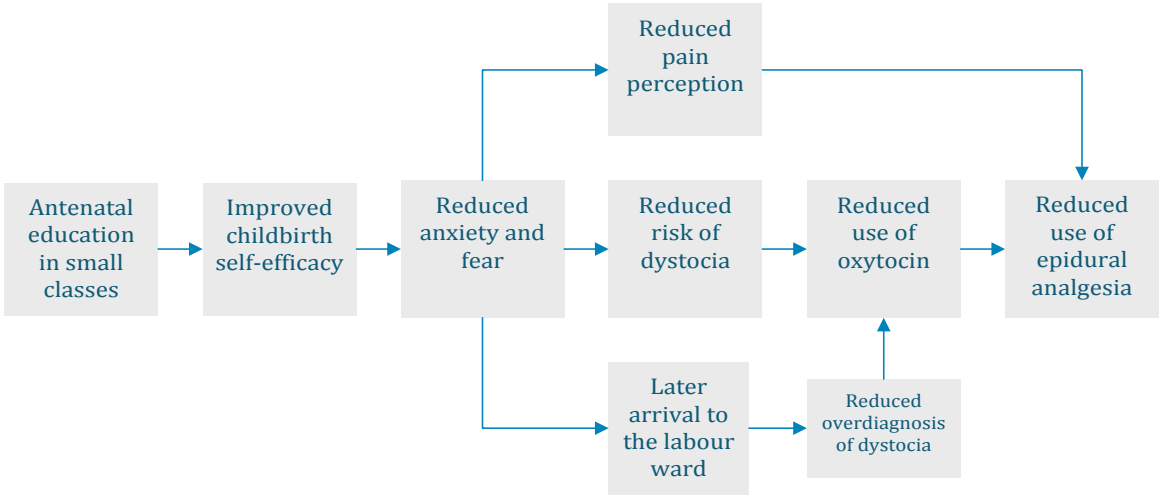
Childbirth self-efficacy reflects a woman's trust in her ability to cope with labour and birth and may influence pain perception, anxiety levels, and obstetric interventions. Clinical studies have found that more than half of the variance in labour pain in the latent phase and more than a third of the variance in pain in the active phase could be explained by the women's confidence in her ability to cope with labour (36, 37). Also, a number of cross-sectional studies have shown an inverse association between childbirth self-efficacy and fear of childbirth (38-40). Self-efficacy is central to the individual's motivation when facing obstacles and aversive experiences (35). Women with low childbirth self-efficacy may therefore have difficulty in generating motivation for coping with the labour experience (39). This may result in psychological withdrawal from the labour experience by use of pain relief or demanding to have elective caesarean section (39). A Swedish study reported that a higher proportion of women with low levels of childbirth self-efficacy used epidural analgesia during labour compared to women with high childbirth self-efficacy, but no difference was seen regarding other obstetric outcomes (40). Also, childbirth self-efficacy may influence the timing of arrival to the labour ward. It has been suggested that women with increased ability to cope with the early phase of labour while being at home will tend to arrive later at the labour ward (41, 42).

2.4. ANTENATAL EDUCATION IN SMALL CLASSES – A WAY TO IMPROVE SELF-EFFICACY AND REDUCE USE OF PAIN RELIEF AND OBSTETRIC INTERVENTIONS

The promotion of self-efficacy beliefs during pregnancy may reduce anxiety, and strategies designated to increase childbirth self-efficacy should include one or more of the four sources of self-efficacy (40). These sources of self-efficacy could be provided through antenatal education. Past experiences of mastering specific situations are the strongest source of self-efficacy according to Bandura (35). However, even primiparous women may benefit from antenatal education in small classes by hearing from other pregnant women's experiences (vicarious experience) and by encouragement from the other participants and a midwife (social persuasion). By these means, antenatal education in small classes may increase the

woman's trust in her ability to cope with early labour and hence reduce anxiety and fear during the early and middle stage of labour. This may in turn reduce the likelihood of early admission, use of pain relief, and obstetric interventions (figure 2.1). This potential of antenatal education classes was examined in a recent Iranian trial that examined the effect of a self-efficacy promoting antenatal education programme in small classes versus no education (43). The trial reported lower fear of childbirth and higher childbirth self-efficacy scores among women in the intervention group (43).

Figure 2.1. Model of the potential pathway from antenatal education in small classes to use of epidural analgesia



In addition to the potential of increasing self-efficacy, antenatal education in small classes may increase the women's knowledge uptake due to the possibility of being actively involved in the learning process (12) in contrast to the auditorium-based lectures which fosters passive learning (12). Thus, antenatal education in small classes with information on positive and negative side effects of pharmacological pain relief as well as general information on the birth process may reduce the use of pain relief and reduce the risk of obstetric interventions.

2.5. Former trials assessing the effect of antenatal education in small classes on pain relief and obstetric interventions

In the Western world, only five randomised trials have examined the effect of attending antenatal education in small groups, and compared the effect of these with the effect of other forms of education on outcomes like use of pain relief or obstetric interventions (41, 44-47). Among these trials conclusions are conflicting (48).

Duffy et al. examined the effect of an extra breastfeeding session on breastfeeding rates but also reported effects of the intervention on obstetric interventions. They found no effect of the intervention on caesarean section or vacuum extraction (47).

A trial by Rouhe et al. conducted in Finland compared the effect of a group-based psycho-educational intervention directed towards women with severe fear of childbirth versus written information in the form of a letter addressing fear of childbirth (44). Rouhe and colleagues found that the intervention significantly increased the likelihood of spontaneous vaginal delivery but found no effect on the use of epidural analgesia, elective or emergency caesarean section, vacuum extraction, or induction of labour (44).

In a Danish trial, Werner et al. compared a self-hypnosis intervention versus standard care (no antenatal education) and reported no effect on use of epidural analgesia, spontaneous delivery, elective caesarean section, vacuum extraction, or oxytocin augmentation. However, they reported an increased risk of emergency caesarean section among participants in the self-hypnosis group (45).

The trials by Rouhe et al. and Werner et al. were performed among women screened positive for fear of childbirth limiting generalisation of results to the general population.

A trial by Downe et al. assessed the effect of a self-hypnosis intervention in small classes versus standard care (no education) among women from a non-high risk population. They found no effect of two 90-minute group sessions on use of epidural analgesia, caesarean section, instrumental delivery, or spontaneous vaginal delivery (46).

One Danish trial by Maimburg et al.; the 'Ready for Child' trial, examined the effect of antenatal education classes in small groups versus standard care (no antenatal education) among women recruited from a diverse population group. This trial reported significantly lower use of the primary trial outcome, epidural analgesia, among women participating in the small classes. There were no significant intervention effects on use of other types of pain relief or obstetric interventions (41).

Due to the sparse evidence from randomised trials, research about the effects of antenatal education in small classes on birth related outcomes is still needed (48). We therefore conducted a randomised trial to examine the effect of a general antenatal education programme in small classes versus standard education carried out as auditorium-based lectures.

3. OBJECTIVES OF THE THESIS

As stated in the introduction, the overall aim of my PhD study was to assess the effect of antenatal education in small classes on use of pain relief, obstetric interventions, and childbirth self-efficacy.

I have the following objectives for this thesis:

1. To assess the current evidence for the effect of antenatal education in small classes versus other types of education (paper I and II).

2. To design a randomised trial examining the effects of antenatal education in small classes compared to auditorium-based lectures on use of epidural analgesia, other types of pain relief, obstetric interventions, and psycho-social outcomes (paper III).

3. To test the validity of the data source used for the obstetric trial outcomes; the Obstetric Database (paper IV).

4. To examine the effect of the NEWBORN trial on use of epidural analgesia, other types of pain relief, and obstetric interventions (paper V).

5. To examine the effect of the NEWBORN trial on the intermediate trial outcome; childbirth self-efficacy (paper VI).

6. To update the evidence for the effect of antenatal education in small classes versus other types of education on obstetric outcomes allowing new trials to be included.

4. OVERVIEW OF PAPERS

The thesis is based on the following papers:

I. **Brixval CS**, Axelsen SF, Andersen SK, Due P, and Koushede V. The effect of antenatal education in small classes on obstetric and psycho-social outcomes: a systematic review and meta-analysis protocol. Systematic Reviews. 2014;3:12.

II. **Brixval CS**, Axelsen SF, Lauemøller SG, Andersen SK, Due P, and Koushede V. The effect of antenatal education in small classes on obstetric and psycho-social outcomes-a systematic review. Systematic Reviews. 2015;4(1):1-9.

III. Koushede V, **Brixval CS**, Axelsen SF, Lindschou J, Winkel P, Maimburg RD, Due P, and the NEWBORN trial group. Group-based antenatal birth and parent preparation for improving birth outcomes and parenting resources: Study protocol for a randomised trial. Sexual & Reproductive Healthcare, 2013;4(3):121-6.

IV. **Brixval CS**, Thygesen LC, Johansen NR, Rørbye C, Weber T, Due P, and Koushede V. Validity of a hospital-based obstetric register using medical records as reference. Clinical Epidemiology, 2015;7:509-15.

V. **Brixval CS**, Thygesen LC, Axelsen SF, Gluud C, Winkel P, Lindschou J, Weber T, Due P, and Koushede V. Effect of antenatal education in small classes versus standard auditorium based lectures on use of pain relief during labour and of obstetric interventions: results from the randomised NEWBORN trial. BMJ Open. 2016;6(6):e010761.

VI. **Brixval CS**, Axelsen SF, Thygesen LC, Due P, and Koushede V. Antenatal education in small classes may increase childbirth self-efficacy: results from a Danish randomised trial. Sexual & Reproductive Healthcare, 2016, in press.

5. METHODS AND RESULTS

The methods and the main findings of the papers I-VI are presented in the sections below. For a more detailed description of the methods and results from the individual papers, please see the papers.

In this thesis I use a variety of methods in the different papers. For the sake of readability, the methods and results sections are combined. Section 5.1 is an overview of the systematic review (objective 1). In this thesis I focus on outcomes related to birth. Section 5.2 describes the design of the NEWBORN trial (objective 2), including recruitment and randomisation, content of intervention and control conditions, data sources, and participant flow and attendance. In section 5.3, I present the examination of the validity of the Obstetric Database that is used for assessment of the obstetric outcomes (objective 3). Section 5.4 is an overview of the statistical methods used and the results of the NEWBORN trial concerning use of pain relief and obstetric outcomes (objective 4). Section 5.5 is an overview of the statistical methods used and the results of the NEWBORN trial on childbirth self-efficacy (objective 5). Lastly, in section 5.6, I present an update of the systematic review including results from the NEWBORN trial (objective 6).

5.1. SYSTEMATIC REVIEW OF THE EFFECT OF ANTENATAL EDUCATION IN SMALL CLASSES (PAPER I AND II).

In order to assess the currently available evidence for the effectiveness of antenatal education in small classes we conducted a systematic review (48). The systematic review was carried out using the Cochrane Handbook for Systematic Reviews of Interventions as a guide (3).

Prior to conducting the literature search, the systematic review was registered in the International Prospective Register of Systematic Reviews (PROSPERO) (registration number CRD42013004319) (49) and the methodology was published as a review protocol (paper I) (50).

MATERIAL AND METHODS

In this section, I give an overview of the methods and main results of the systematic review. For detailed descriptions of the methods and results, including the specific search strategy, risk of bias assessment, tables with characteristics of included and excluded trials, effect tables, and forest plots please see the additional files for paper II.

Eligibility criteria

Preparation for birth and parenthood is sensitive to culture and contextual factors, such as the organisation of the healthcare system. The purpose of the systematic review was to guide decision-making in Western

countries, and therefore trials conducted in developing countries were excluded and only trials conducted in the Western world – defined as OECD membership countries (51) – were included. Trials that compared antenatal education programmes in small classes with no or other types of antenatal education programmes were included. Trials that compared two kinds of antenatal education in small classes were excluded due to the difficulty of assessing the effect of antenatal education in small classes as an experimental condition as only the content varied between the experimental and the control condition.

Trials with co-interventions, such as extra individual sessions, e.g. extra individual consultations with the midwife or home visits provided by e.g. a midwife or health visitor, were excluded as these co-interventions might have influenced the effect of the intervention beyond the effect of the classes. We accepted extra written material to the experimental group. In trials where co-interventions were delivered equally to both groups, the presence of co-interventions were accepted and not a reason for exclusion of the trials. In trials where the intervention was 'boosted' by later individual consultations, we used the measurement shortly before the individual consultation to consider the effect. In cases where the content of the experimental or control condition was unclear or information incomplete, we contacted the first author by e-mail. We contacted 19 authors and received replies from six of these.

Literature search

The search was performed March 5th 2014, and the following databases were sought without language or publication date restriction; Medline, EMBASE, CENTRAL, CINAHL, Web of Science, and PsycINFO. In addition to published trials, we searched for relevant trials in citations from identified papers and former reviews, and contacted authors of the included trials for information on unpublished results in those cases where a valid e-mail address could be found.

Outcome measures

Due to the many and comprehensive aims of antenatal education we chose to assess the effect of antenatal education in small classes on a broad range of outcomes. We assessed psycho-social outcomes, e.g. antenatal and postnatal depression, anxiety, and satisfaction with relationship as well as outcomes related to labour and birth, e.g. pain relief and obstetric interventions.

The primary outcomes examined were: use of pain relief during labour, obstetric interventions, psychological and social adjustment to parenthood, and antenatal and postnatal depression and anxiety. Secondary outcomes were: knowledge acquisition, maternal sense of control/active decision-making during labour and birth, partner involvement at birth, breast feeding success, infant care abilities, social support, relationship satisfaction, and divorce or separation. In this thesis, I focus on presenting the results regarding childbirth self-efficacy, pain relief, and obstetric interventions.

Trial selection and data extraction

Trials were selected by two independent assessors in two steps – first by screening of titles and abstracts and next, through assessment of full-text papers. Data from the included trials was extracted to summary tables containing information on the following: study design, inclusion and exclusion criteria, description of the experimental and control conditions, and outcomes of interest to the review.

Risk of bias assessment

Separate domains

Risk of bias in each trial was assessed by two independent assessors using the approach recommended by the Cochrane Collaboration (3) and included the following domains: randomisation sequence generation; allocation concealment; blinding of participants, educators, and outcome assessors; incomplete outcome data; selective outcome reporting; and other sources of bias.

Each trial was evaluated according to each of the above-mentioned bias domains as either 'low risk of bias', 'unclear ', or 'high risk of bias'.

Blinding of participants and educators is often not possible in these types of trials. Subjectively measured outcomes are more prone to bias related to blinding, as the participant him or herself is the outcome assessor. Objectively measured outcomes can be assessed blinded although the participants and educators are not blinded. In line with the recommendations from the Cochrane Collaboration (3), we therefore assessed risk of bias related to blinding separately for subjective and objective outcomes.

Similarly, bias related to incomplete outcome data may also be dependent of the outcome under study. We therefore assessed risk of bias related to incomplete outcome data separately for objective and subjective outcomes.

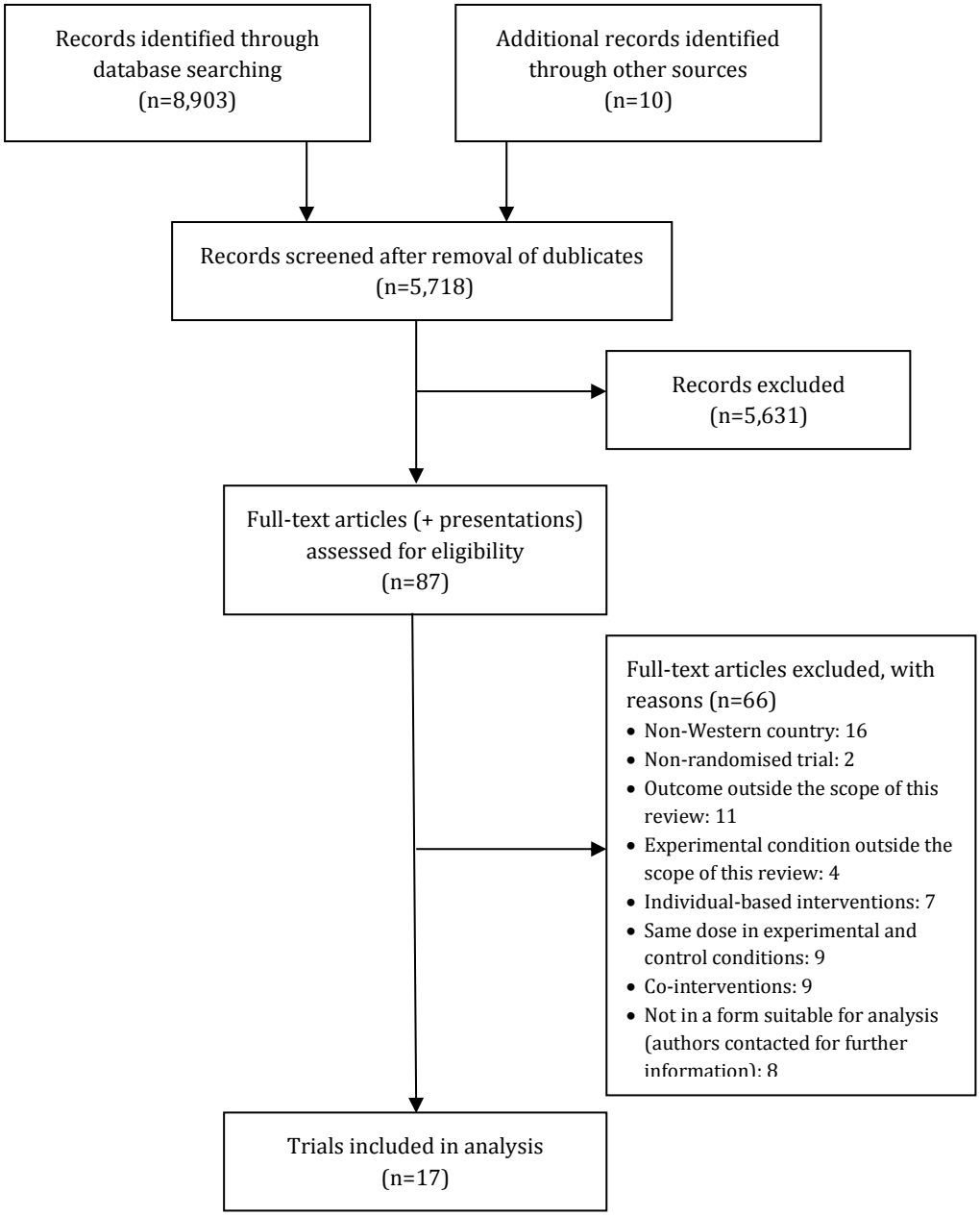
Overall risk of bias

We assessed the overall risk of bias for each trial. Trials were rated as overall 'low risk of bias' if the trial scored 'low risk of bias' in all the six separate bias domains. Otherwise, the trial was scored 'high risk of bias'. Due to the inherent nature of the ability to blind participants and educators, we rated the trial as 'moderate risk of bias' if all trial bias domains were rated as 'low risk of bias' with the exception of blinding of participants and educators. The rating of overall risk of bias was also done separately for subjective and objective outcomes.

SUMMARY OF RESULTS

A flow diagram of the trial selection process is shown in figure 5.1. In total, 17 trials were included (41, 44, 45, 47, 52-64). The trials included a total of 6,507 randomised women and 961 men, with a range from 74 to 1,193 participants per trial.

Figure 5.1. Flow diagram of trial selection



There were large variations in form as well as content of the experimental and control conditions between the trials. The dose of the interventions varied from a single one-hour session to 24 sessions each lasting 2.5 hours. Also, the population varied greatly across trials; some trials focused on prevention of a specific condition among participants at specific risk, e.g. women with severe fear of child birth, whereas other trials targeted a broader population group. After inspecting the included trials, we decided to omit the pre-planned meta-analyses from the systematic review. Meta-analyses were no longer considered appropriate

due to the large heterogeneity of experimental and control condition as well as trial population across trials.

In conclusion, based on the findings from the systematic review it was not possible to draw definitive conclusions on the effect of antenatal education in small classes on obstetric and psycho-social outcomes. Two of the primary outcomes are of relevance to this specific thesis; use of pain relief and obstetric interventions. Use of pain relief was reported in three trials (41, 44, 45) and obstetric interventions were reported in four trials (41, 44, 45, 47). We found no trials reporting on the effect of antenatal education in small classes on one of the secondary outcomes; childbirth self-efficacy. A complete review of all effect estimates from all trials included in the systematic review can be found in the additional files for paper II. Below I give an overview of the results from the trials examining the effect on use of pain relief and obstetric interventions.

Pain relief

Three trials examined the effect of antenatal education classes on at least one type of pain relief (41, 44, 45). Maimburg et al. assessed the effect of a group-based antenatal training programme compared with standard care on a range of both pharmacological and non-pharmacological pain relief outcomes among 1,193 women (41). They found no statistically significant differences between the intervention and control group on use of pain relief, except for a small but statistically significant protective effect on use of epidural analgesia (relative risk=0.84, 95% CI: 0.73-0.98). Relative risks for the other types of pain relief varied between 0.68 and 1.27 (41). Werner et al. compared a self-hypnosis intervention with standard care among 727 women and found no effect on use of epidural analgesia as pain relief during labour (relative risk=1.04 (95% CI: 0.82-1.32)) (45). No other types of pain relief were examined in this trial. Rouhe et al. examined the effect of a group-based psycho-educative intervention directed towards women with severe fear of childbirth among 371 women. The control group received a letter with a recommendation to discuss their fear of childbirth (44). The authors examined the effect of the intervention on use of epidural analgesia but not on any other types of pain relief. The relative risk of receiving epidural analgesia was 1.13 (95% CI: 0.96-1.33) (44). Across the three trials, the relative risks were close to 1.00 and confidence intervals were generally narrow indicating that the observed results were not merely reflecting inadequate sample sizes.

Obstetric interventions

Four trials reported effects of antenatal classes on obstetric interventions in the birth process (41, 44, 45, 47). Overall, there were no effects on labour induction, elective caesarean section, vacuum extraction, forceps, and oxytocin augmentation. Two trials reported significant effects on spontaneous vaginal deliveries and emergency caesarean section (44, 45). Rouhe et al. found that a group-based psycho-educative intervention directed towards women with severe fear of childbirth increased the likelihood of spontaneous vaginal deliveries (relative risk=1.33, 95% CI: 1.11-1.61). In addition, they reported a

borderline significant protective effect on overall caesarean section (relative risk=0.70, 95% CI: 0.49-1.01) (44). Werner et al. found that a self-hypnosis intervention increased the risk of having emergency caesarean section (relative risk=1.52, 95% CI: 1.02-2.27) (45). Maimburg et al. found no significant effects of a general antenatal training program on any of the registered obstetric interventions (41). In these three trials, the relative risks for obstetric interventions ranged from 0.51 to 1.52, and confidence intervals were generally narrow, indicating that the observed results were not merely reflecting inadequate sample sizes. Duffy et al. examined the effect of an antenatal group teaching session aimed at increasing breast feeding prevalence among 73 women and reported on obstetric interventions. They found no effect of the intervention on vaginal delivery, caesarean section, vacuum extraction or forceps (47). The relative risks reported in this trial ranged from 0.67 to 1.00 with wide confidence intervals, reflecting lack of effect but also lack of power in the trial due to the small sample size and small number of events.

At the end of this chapter, I present updated analyses of the systematic review including new trials (section 5.6).

5.2. Design of the NEWBORN trial (paper III)

Overview and design

The overall aim of the randomised NEWBORN trial was to examine the effect of a structured antenatal education programme in small classes versus standard care, carried out as auditorium-based lectures, on outcomes related to birth and parenthood.

The NEWBORN trial was conducted from 2012 to 2014 at the largest birth site in Denmark: Hvidovre Hospital situated in the Copenhagen Capital Region. At Hvidovre Hospital, more than 6,500 deliveries take place every year and the hospital's catchment area comprises a diverse population regarding socio-demographic characteristics.

Prior to enrolment of the first participant, the trial was registered at ClinicalTrials.gov (identifier: NCT01672437), and a trial protocol was published describing among other things the trial design, participant eligibility criteria, outcome measures, and the plan for the statistical analyses (paper III) (65). In addition, a more detailed trial protocol was developed and placed at the trial web page (http://www.interventionsforskning.dk/newborn-project-protocol/) (66).

We performed a sample size calculation based on the primary outcome of the trial; use of epidural analgesia. In addition, we performed power calculations on the secondary outcomes, i.e. perceived stress, parenting stress, and parenting alliance to ensure adequate power in these outcomes as well. The calculations were based on data from a previous Danish trial (41) and data from the Obstetric Database at

Hvidovre Hospital. Our original sample size calculations were based on a power of 0.90 and a significance level of 0.05 requiring randomisation of 2,350 women.

The Danish Ethics Committee for the Capital Region of Denmark reviewed the trial protocol and concluded that formal ethical approval was not required according to Danish legislation as no human physiological interventions were conducted (protocol number H-4-2012-FSP). The trial was registered and listed in the Danish Data Protection Agency (reference number 2011-54-1289).

PARTICIPANTS AND RANDOMISATION

Women who were ≥18 years old, pregnant with a single child, due to give birth at Hvidovre Hospital, and able to speak and understand Danish were eligible to participate in the trial. All women fulfilling these inclusion criteria were invited to participate in the NEWBORN trial via a written, mailed invitation prior to their first visit to the midwife. Initially, only primiparous women were eligible for participation, but due to slow recruitment also multiparous women were included approximately six months into the recruitment period in order to ensure adequate statistical power (66).

Included in the invitation was an instruction on how to fill in a web-based baseline questionnaire. After completing the questionnaire and mailing a signed consent form to the trial research group, participants were randomised according to a computer-generated allocation sequence of 1:1 to the intervention or control group by a research assistant. The allocation was stratified for parity (primiparous or multiparous) and vulnerability (yes or no as defined by their general practitioner at the first pregnancy consultation in gestation week 6-10). In Denmark, the health authorities have specified groups of women in need for special attention – here named vulnerable women. The assessment of vulnerability in pregnant women is performed as a standard procedure in order to assess each woman's need for extra care during the pregnancy (7). At Hvidovre Hospital there are eight criteria listed for vulnerability, for example; former or current psychiatric disorder, adverse psycho-social background, or concerns about parenting skills. The general practitioner categorised the women as vulnerable if she met one or more of these criteria. For more information on the criteria, please see appendix 12.1. The stratification was conducted to prevent imbalances in these assumed important prognostic factors for use of epidural analgesia, i.e. vulnerability and parity.

INTERVENTION GROUP

The development of the NEWBORN programme was guided by an Intervention Mapping approach (67) and designed to meet the recommendations from the Capital Regions' Birth Planning Committee and the Danish Health Authority regarding form and content (7). First, we established a planning group consisting of different stakeholders, e.g. midwives, health visitors, family therapists, and representatives from the interest organisation *Parenting and Childbirth.* The planning group delivered inputs for the form and content of the programme. Next, a working group developed the content of each of the four sessions. Each

of the first three sessions was pilot tested among pregnant women and their partners in small groups with participation of six to eight couples in each group. A midwife with experience in teaching facilitated these classes. After the pilot test, the material was adjusted according to the lessons learned.

The overall aim of the programme was to strengthen couple relationships and improve information and problem solving skills for expectant parents in order to ease birth and the transition to parenthood. As recommended by learning theorists (12) and in line with women's own preferences (13), the education was organised in small groups.

Groups of six to eight women and their partners met three times during pregnancy (gestation week 25, 33, and 35) and one time five weeks after the expected due date, for the duration of 2.5 hours per session. The topics of the sessions were: couple relationship and communication strategies; labour and birth; infant nutrition, care for the infant, and symptoms of postnatal depression; and birth experience and the first period with the newborn (figure 5.2). Throughout the programme there was a focus on increasing self-efficacy in relation to the different topics touched upon, e.g. childbirth self-efficacy. Also, the programme aimed at enhancing social network among the participants and highlighted the importance of partner support.

The sessions were facilitated by a midwife who followed a detailed teaching manual developed for the trial. During the sessions the midwife held short presentations on relevant topics. A large part of the sessions consisted of other types of activities, such as discussions with own partner, group work, and plenary discussions. These elements were integrated to stimulate learning by engaging the participants.

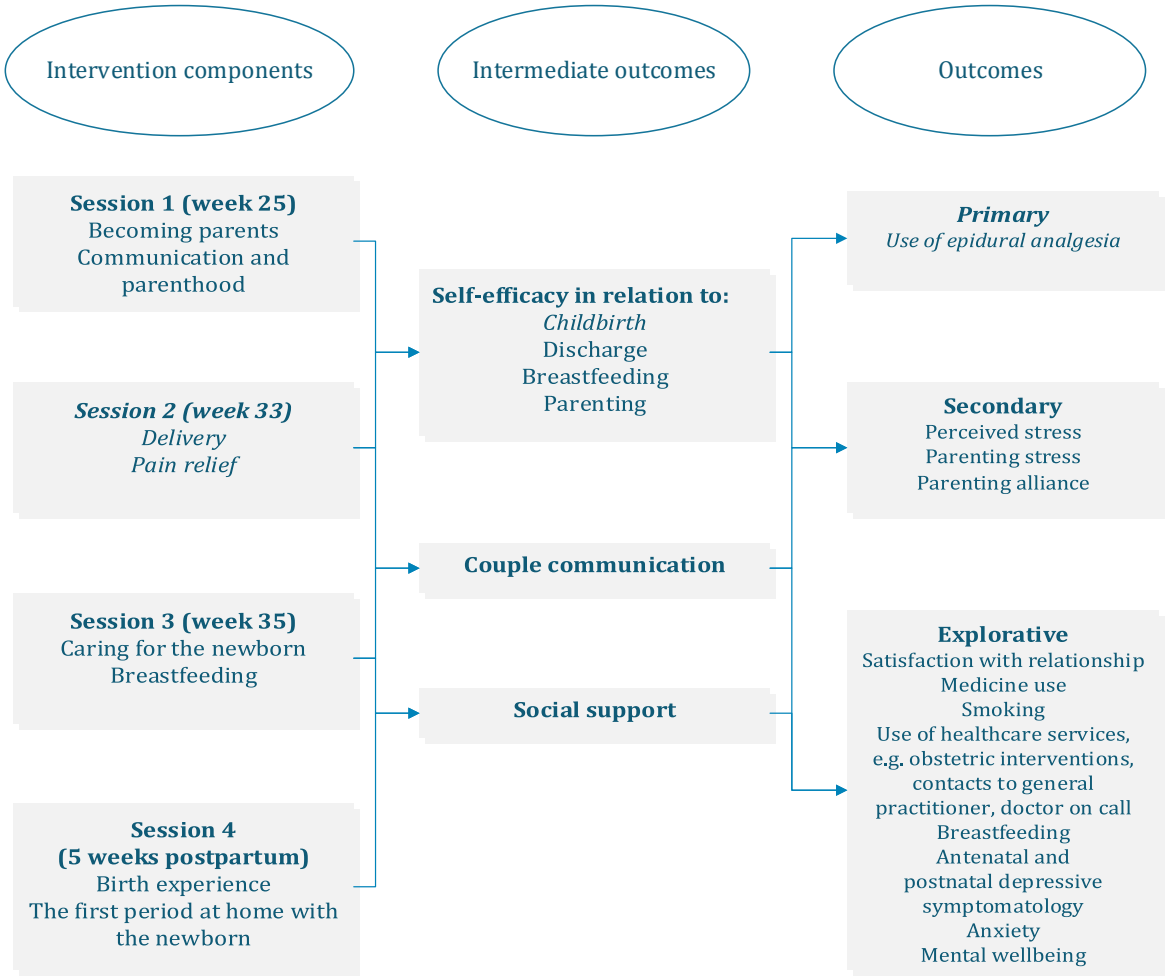
The session in the 33rd week of gestation focused on pain relief and the birth process. The aim of this session was to enhance the participant's existing knowledge and understanding of the normal course of labour, pain relief, and what might be expected if an obstetric intervention became necessary. During the session, the women were asked to discuss their thoughts and previous experiences with coping with pain and physical and mental strain, and to consider whether they thought they could use any of these strategies during labour. Next, the participants discussed their thoughts and knowledge on various methods of pain relief. Plenary discussions and summing up thoughts and ideas were used so that participants could learn from and be inspired by one another (65).

In addition to the classes, the participants had access to a patient-network website specifically developed for the trial. At this website the participants could gain further information on topics presented at the sessions and watch small film clips. For example, there was film clips of a midwife going through different types of pain relief and obstetric interventions and giving advice about what to do at home in the early phase of labour before going to the labour ward. In addition, the webpage functioned as a platform for communication with the other participants and gave the opportunity to consult online with a midwife and a

health visitor. Prior to each session the participants were encouraged by the midwife to use the website as preparation for the session.

Twenty five midwives with varying professional seniority and teaching experience facilitated the sessions. Midwives were recruited through job advertisements in a newsletter for the staff at Hvidovre Hospital and by informing about the trial at staff meetings. The first 15 midwives who signed up for teaching attended a one-day workshop prior to the start of the intervention. Due to a large dropout of midwives, a second half-day workshop was arranged for midwives signing up for teaching after intervention start. Towards the end of the intervention period, new midwives who signed up for teaching attended a session held by an experienced facilitator, but due to time constraints they did not receive a workshop.

Figure 5.2. Programme logic model of the NEWBORN trial



CONTROL GROUP

Women in the control group received the existing standard care offer from Hvidovre Hospital consisting of two antenatal lectures of two hours each on the issues delivery and breastfeeding in an auditorium with participation of up to 250 people. To avoid contamination of conditions midwives facilitating the NEWBORN programme were not allowed to teach the antenatal lectures in the control group.

Participants in the intervention group as well as in the control group were permitted to make use of concomitant birth and parent education.

OUTCOMES

The primary trial outcome was use of epidural analgesia during labour. Secondary outcomes were perceived stress, parenting stress, and parenting alliance. Exploratory outcomes included amongst others antenatal and postnatal depressive symptomatology, relationship satisfaction, breastfeeding, and obstetric interventions. All outcomes were related to the components in the intervention and assumed to be amendable to change through the intermediate outcomes; self-efficacy, couple communication, and social support (figure 5.2).

In this PhD thesis, I examine the effects of the intervention on outcomes directly related to birth, i.e. use of epidural analgesia, other types of pain relief, obstetric interventions, and the intermediate trial outcome; childbirth self-efficacy.

BLINDING

It was not possible to blind participants or teaching midwives in the NEWBORN trial. The outcome assessors of the obstetric outcomes; midwives and physicians at the labour ward were not informed about the women's participation in the trial. Further, in all communication from the trial, the NEWBORN trial was referred to as a general programme aiming to increase resources for birth and parenthood among expectant parents. The health personnel were not informed that the primary outcome of the trial was use of pain relief. Data were blinded by a data manager and I was, in my role as data analyst, therefore blinded to participants' allocation group during analyses. After completing the statistical analyses and presenting the blinded results for the steering committee, conclusions on trial effects were drawn by the committee. After conclusions were agreed upon, the randomisation code was revealed.

DATA SOURCES AND VARIABLES

Questionnaire data

Data was collected by web-based questionnaires from both parents at five occasions; at baseline, at 37 weeks of gestation, nine weeks after expected due date, six months after expected due date, and one year after expected due date. Data on attendance the pregnant women's participation in the sessions as well as the participants' assessment of the facilitator's adherence to the programme content was assessed by tablet-based questionnaires after each session.

For a thorough description of all variables used in the thesis, please see appendix 12.1. I used one outcome measured by questionnaire in the analyses for this thesis; childbirth self-efficacy. Data on this intermediate trial outcome was collected by the mother's questionnaire in gestation week 37 and was measured by three single items developed for the NEWBORN trial: 1) I believe that I will feel confident at home once labour has begun (e.g. before going to the labour ward), 2) I believe that I can contribute to making the birth a good experience, and 3) I believe that I will be able to handle the birth process no matter how it turns out. All items had the following response categories: totally agree, agree, neither/nor, disagree, and totally disagree. I trichotomised the responses in the following categories: high self-efficacy (totally agree, agree), neither/nor, and low self-efficacy (disagree, totally disagree). Self-efficacy is used as an overarching concept that covers the individuals' confidence in own ability to cope with certain behaviours and the three questions are considered expressions for the woman's confidence in her own ability to cope with the specific situations.

Register data

Data on use of pain relief and obstetric interventions was collected from the hospital-based clinical register at Hvidovre Hospital: the Obstetric Database. During and after labour, midwives register the neonatal and obstetric data, including pharmacological pain relief and obstetric interventions, in the Obstetric Database by ticking an electronic list or adding a diagnosis/treatment according to ICD 10. Postpartum, a specialist doctor or senior midwife goes through every file and adds left out information and supplies specialist diagnoses based on information from the medical records (68).

BASELINE CHARACTERISTICS, PARTICIPANT FLOW AND ATTENDANCE

Recruitment of participants started in August 2012. We expected that about half of all pregnant women would agree to participate and that the required sample size could be reached within one year. However, it turned out to be more difficult to recruit participants than expected, and a revised sample size calculation was therefore performed. Reducing power from 0.90 to 0.80 resulted in a reduction in required sample size from 2,350 women to 1,756 women. The adjustment was done after inclusion of 1,050 participants without inspection of the data (66, 69).

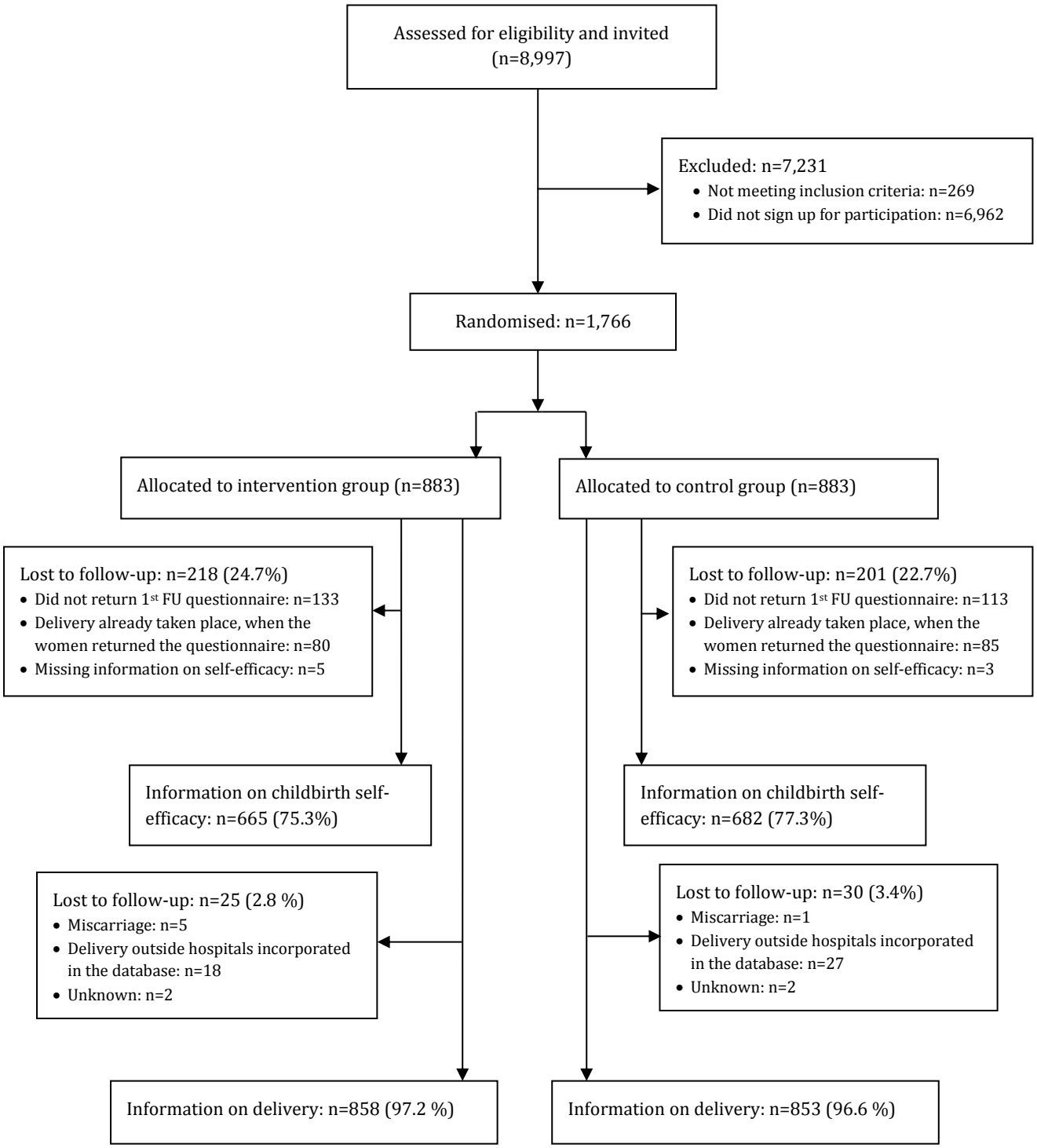
From August 2012 to May 2014; 8,997 women were invited to participate in the NEWBORN trial. Of these, 1,766 women (19.6%) accepted participation and were randomised – 883 women to the intervention group and 883 to the control group (figure 5.3). At baseline, the characteristics among the intervention and control groups were well balanced (table 5.1).

Table 5.1. Baseline characteristics of women enrolled in the NEWBORN trial (n=1,766).

	Intervention (n=883)	Control (n=883)
Age at birth in years (mean (SD))*	30.7 (4.1)	30.8 (4.1)
Nulliparous % (n)	89.1 (787)	88.9 (785)
Vulnerable women % (n)	4.8 (42)	4.8 (42)
Educational level (higher tertiary education) % (n)	75.6 (659)	76.5 (663)
Body Mass Index kg/m² (mean (SD))*	23.4 (4.0)	23.3 (4.1)
Living with child's father (yes) % (n)	93.8 (828)	96.0 (848)
Planned pregnancy (yes/partly) % (n)	90.9 (801)	91.5 (808)
Self-rated physical health status (excellent/very good) % (n)	68.6 (605)	71.2 (628)
Self-rated mental health status (excellent/very good) % (n)	72.0 (635)	75.9 (669)
Not feeling stressed % (n)	48.2 (425)	49.2 (433)
Edinburgh Postnatal Depression Scale score ≥13 % (n)	4.8 (42)	3.2 (28)
Perceived Stress Scale score (mean (SD))	12.5 (5.2)	12.2 (5.2)

* Based on women with birth data (n=1,711).

Figure 5.3. Flow diagram of recruitment, randomisation, and participation in the NEWBORN trial.



The attrition seemed similar in the two groups. Information on childbirth self-efficacy was provided by 75.6% of the women in the intervention group and 77.3% of the women in the control group. A total of 97.2% of the women in the intervention group and 96.6% in the control group had information about their delivery in the Obstetric Database.

ADHERENCE

Adherence to the intervention

In the intervention group, 94% attended session 1, 92% attended session 2, and 82% attended session 3. A total of 69% attended session 4. In this thesis, I focus on attendance before birth, and hence attendance in session 4 is not used in the overall measure of attendance. A total of 73% attended all three sessions prior to delivery. Adherence to the intervention was defined as attendance in all sessions prior to delivery and using the website – 68% adhered to the intervention. In the control group, 71% attended the auditorium-based lecture on birth, 63% attended the lecture on breastfeeding, and 59% attended both lectures.

Adherence to the programme content in session 2

The session most directly related to this thesis is session 2. This session focused on pain relief and the birth process and the overall aim was to enhance the participant's existing knowledge and understanding of the normal course of labour, pain relief, and what might be expected if an obstetric intervention became necessary.

Generally, the facilitator's adherence to the programme content in session 2 was reported high by the participants. More than 97% of the participants reported to have heard about the topics: 'expectations in relation to birth', 'what to do at home in the early phase of labour', 'the normal course of labour, 'pain relief and coping strategies', and 'partner support during labour'. A total of 88% of the participants reported having been through the topic 'when there is a need to intervene in labour'.

5.3. ASSESSMENT OF COMPLETENESS AND VALIDITY OF INFORMATION IN THE OBSTETRIC DATABASE (PAPER IV)

Reliable use of data from registers requires valid and complete data sources. To investigate the reliability and quality of information in the Obstetric Database we conducted a validation study assessing completeness of the database as well as validity of selected indicators (68).

MATERIAL AND METHODS

Completeness of the Obstetric Database was examined by linking data from all women registered in the database as having given birth in 2013 to the National Patient Register. Validity of 11 selected indicators from the database was assessed using medical records as the golden standard. Five of these indicators

concern obstetric interventions and are of relevance to this thesis: use of epidural analgesia, use of oxytocin due to dystocia, vacuum extraction, elective caesarean section, and emergency caesarean section.

Procedure

We expected a positive predictive value of 95% of the information in the Obstetric Database and wanted to estimate this with a confidence interval of 92%–98%. In order to fulfil this, a sample of 203 deliveries was required. To take incomplete data into account we selected a random sample of 250 deliveries in 2013 from the Obstetric Database and retrieved the corresponding electronic medical records. Three deliveries were excluded due to transfers to other hospitals during labour (n=2) and due to missing data on all indicators for unknown reasons (n=1). There is a risk of recording error in the process of the manual inspection of the medical records. We, therefore, chose that two assessors independently reviewed the medical records and registered the indicators in separate documents.

Statistical analysis

To assess the validity of the information in the Obstetric Database, I calculated sensitivity, specificity, positive and negative predictive values as well as proportion of agreement with exact 95% binomial proportion confidence intervals for each of the indicators.

I assessed the inter-rater reliability between the two assessors by calculating the proportion of agreement and Cohen's Kappa coefficient. The Landis and Koch's scale (70) was used to categorise strength of agreement from the Kappa coefficients.

I performed the statistical analyses using SAS v. 9.3, SAS Institute Inc.

SUMMARY OF RESULTS

I found 100% completeness of the Obstetric Database when compared to the Danish National Patient Register. Also, the validity of the Obstetric Database was generally high; the proportion of agreement between the Obstetric Database and medical records ranged from 91.1% to 99.6% for the five indicators relevant for this PhD thesis (table 5.2). Measures of sensitivity, specificity, and positive and negative predictive values ranged from 0.70 to 1.00 indicating high validity of the Obstetric Database (table 5.3).

The proportion of false positive registrations for the indicator 'oxytocin due to dystocia' was high due to registration practices. The clinical practice for registering administration of oxytocin implies that both administrations of oxytocin due to dystocia as well as oxytocin used as part of induction of labour is registered under the same code. For this reason I chose not to use 'oxytocin due to dystocia' as an outcome in the NEWBORN trial as the programme was assumed to influence administration of oxytocin due to dystocia, but it was not expected to influence other use of oxytocin as related to induction of labour.

The inter-rater reliability for the five indicators was generally high, and the proportion of agreement between the two assessors ranged from 94.3% (oxytocin due to dystocia) to 100% for epidural analgesia. Kappa coefficients ranged from 0.83 for oxytocin due to dystocia to 1.00 for epidural analgesia, and are therefore considered at least "almost perfect" according to the Landis and Koch categorisation (data not shown, please see paper IV for details).

Table 5.2. Number of registrations in the Obstetric Database and in medical records and the proportion of agreement (%) (95% confidence interval) for each indicator

Indicator	In the Obstetric Database		Not in the Obstetric Database		Proportion of agreement (%)
	In medical records	*Not in medical records*	*In medical records*	*Not in medical records*	
Epidural analgesia	68	0	3	176	98.8 (96.5-99.8)
Oxytocin due to dystocia	46	20	2	179	91.1 (86.8-94.3)
Vacuum extraction	21	0	2	224	99.2 (97.1-99.9)
Elective caesarean delivery	23	1	1	222	99.2 (97.1-100.0)
Emergency caesarean delivery	35	0	1	211	99.6 (97.8-100.0)

Table 5.3. Sensitivity, specificity, positive and negative predictive values (95% confidence interval) for each indicator

Indicator	Sensitivity	Specificity	Positive predictive value	Negative predictive value
Epidural analgesia	0.96 (0.88-0.99)	1.00 (0.98-1.00)	1.00 (0.95-1.00)	0.98 (0.95-1.00)
Oxytocin due to dystocia	0.96 (0.86-0.99)	0.90 (0.85-0.94)	0.70 (0.57-0.80)	0.99 (0.96-1.00)
Vacuum extraction	0.91 (0.72-0.99)	1.00 (0.98-1.00)	1.00 (0.84-1.00)	0.99 (0.97-1.00)
Elective caesarean delivery	0.96 (0.79-1.00)	1.00 (0.98-1.00)	0.96 (0.79-1.00)	1.00 (0.98-1.00)
Emergency caesarean delivery	0.97 (0.85-1.00)	1.00 (0.98-1.00)	1.00 (0.90-1.00)	1.00 (0.97-1.00)

5.4. EFFECT OF THE NEWBORN TRIAL ON USE OF PAIN RELIEF AND OBSTETRIC INTERVENTIONS (PAPER V)

STATISTICAL ANALYSES

Main analyses

I tested differences in frequency of use of epidural analgesia, other types of pain relief, and obstetric interventions between the intervention and control group in logistic regression models adjusted for the stratification variables; parity and vulnerability. Odds ratios (OR) and 95% confidence intervals (95% CI) were calculated. This primary analysis was based on participants with information on the outcome (modified intention-to-treat analysis).

Missing data

The proportion of missing data on use of epidural analgesia was 3%. I assessed the possible impact of missing data on epidural analgesia in two ways: 1) I tested whether missing values were missing completely at random (MCAR) by Little's test (71). By this method the assumption that no identifiable pattern exists to the missing data is tested. 2) I conducted a 'worst case' and a 'best case' scenario sensitivity analysis, which is considered an appropriate sensitivity analysis when a binary outcome is missing in only a small proportion of participants (72, 73). In the worst case scenario missing values of epidural analgesia in the intervention group were imputed by a "yes" and missing values of the control group were imputed by a "no". In the best case scenario missing values of epidural analgesia in the intervention group were imputed by a "no" and missing values of the control group were imputed by a "yes". The idea behind the best-worst case analysis is to examine the results' robustness to missing data in the most extreme form.

Sensitivity analyses

To further investigate the impact of attending other types of antenatal education and adherence to the intervention the following additional post-hoc sensitivity analyses were performed:

1. To assess the impact of the participant's use of concomitant birth and parent preparation, I used logistic regression analyses to calculate adjusted odds ratios for epidural analgesia excluding participants who used concomitant preparation.
2. To take into account the adherence to the intervention, I used logistic regression analysis to calculate adjusted odds ratios for use of epidural analgesia between the intervention and control group in two per protocol analyses: 1) participants in the intervention group who adhered to the intervention versus all participants in the control group; 2) participants in the intervention group who adhered to the intervention versus participants in the control group who participated in both antenatal lectures. Adherence to the intervention was defined as attending all three sessions before

birth and using the website. Only the first three sessions were relevant for the obstetric outcomes as the fourth session was held after delivery.

I performed the statistical analyses using SAS v. 9.3, SAS Institute Inc.

SUMMARY OF RESULTS

A total of 97.2% of the women in the intervention group and 96.6% in the control group had information about childbirth in the Obstetric Database. The proportion of missing data on use of epidural analgesia were hence below 5% and Little's test for MCAR was insignificant (p=0.64). Therefore, no imputation of missing values on use of epidural analgesia was performed as described in the statistical analysis plan (65). The modified intention-to-treat analysis therefore included 1,711 participants.

Main analyses

The analyses showed no statistically significant effect of the NEWBORN intervention on pain relief and obstetric interventions. Among women in the intervention group, 30.5% received epidural analgesia compared with 29.1% in the control group (adjusted OR=1.10 (95% CI: 0.87-1.34)). Also, there were no statistically significant differences between the two groups regarding other types of pain relief or obstetric interventions (table 5.4).

Table 5.4. Adjusted odds ratios (OR) (95% confidence interval (CI)) for use of pain relief and obstetric interventions when using the control group as reference#.

	Intervention n=858	Control n=853	Adjusted OR (95% CI)*
Pain relief			
Epidural analgesia % (n)	30.9% (265)	29.1% (248)	1.10 (0.87-1.34)
Pudendal nerve block % (n)	9.2% (79)	7.5% (64)	1.25 (0.89-1.77)
Water immersion % (n)	18.3% (157)	17.4% (148)	1.07 (0.83-1.37)
Acupuncture % (n)	13.4% (115)	13.6% (116)	0.98 (0.74-1.30)
Intracutaneous sterile water injection % (n)	8.6% (74)	9.4% (80)	0.91 (0.65-1.27)
Morphine % (n)	7.2% (62)	5.6% (48)	1.31 (0.89-1.94)
Nitrous oxide % (n)	0.5% (4)	0.9% (8)	0.50 (0.15-1.66)
Obstetric interventions			
Vacuum extraction % (n)	15.4% (132)	14.9% (127)	1.04 (0.80-1.36)
Emergency caesarean section % (n)	17.4% (149)	17.2% (147)	1.01 (0.78-1.30)
Elective caesarean section % (n)	4.0% (34)	4.9% (42)	0.80 (0.50-1.27)

Analyses are based on the modified intention-to-treat population (n=1,711).
* Adjusted for parity and vulnerability.

Best-worst case analysis

Results from the best case scenario showed no statistically significant difference between the intervention group and control group on use of epidural analgesia (adjusted OR=0.93 (95% CI: 0.76-1.14)). In the worst case scenario, the results indicated a negative impact of the intervention (adjusted OR=1.25 (95% CI: 1.02-1.54)), i.e. that more women in the intervention group used epidural analgesia compared with women in the control group. The results indicate that the effect of the intervention lies between the two scenarios; i.e. between an OR of 0.93 (95% CI: 0.76-1.14) and an OR of 1.25 (95% CI: 1.02-1.54).

Sensitivity analyses

Use of private antenatal education was considerably higher among participants in the control group (38.7%) than among participants in the intervention group (25.0%). Results from analyses excluding women who made use of concomitant birth and parent preparation were similar to the results from the modified intention-to-treat analysis i.e. there were no statistically significant difference between the two groups.

Also, results from per-protocol analyses comparing use of epidural analgesia among participants adhering to the intervention with the control group were consistent with the results from the modified intention-to-treat analysis (data not shown, please see paper V for details).

5.5. Effect of the NEWBORN trial on childbirth self-efficacy (paper VI)

In order to further investigate the mechanism between antenatal education in small classes and use of pain relief and obstetric interventions as described in figure 2.1, I examined the effect of the NEWBORN trial on the intermediate trial outcome; childbirth self-efficacy (74).

Statistical analyses

Main analyses

Childbirth self-efficacy was measured by three items concerning the woman's confidence in her own ability to cope with certain situations. I tested differences in childbirth self-efficacy between the intervention group and control group in multinomial logistic regression models adjusted for the protocol specified stratification variables; vulnerability and parity. Multinomial logistic regression models allow for assessment of the intervention effects of an outcome variable of more than two levels (75). Odds ratios (OR) and 95% confidence intervals (95% CI) were calculated. High childbirth self-efficacy was used as reference category, and I therefore calculated the odds for having low self-efficacy versus high self-efficacy in the intervention group divided by the odds for having low self-efficacy versus high self-efficacy in the control group. Likewise, I calculated the odds for having "neither/nor" self-efficacy versus high self-efficacy in the intervention group divided by the odds for having "neither/nor" self-efficacy versus high self-efficacy

in the control group. This was done for each of the three childbirth self-efficacy measures in three separate multinomial logistic regression models.

Information on childbirth self-efficacy was provided by 75.6% of the women in the intervention group and 77.3% of the women in the control group. I therefore used inverse probability weighting to account for the potential bias related to the missing values (76). In this method, participants with full response are weighted by the inverse of their probability of being a full respondent to ensure generalisability to the full trial sample. This means that only participants with full response are analysed directly but weights are assigned to the respondents with full response based on the estimated probabilities of being a full respondent (77).

I used the following variables derived from the baseline questionnaire to estimate the weights: self-rated physical health, self-rated mental health, feeling stressed, and occupational status. For a detailed description of the variables, please see appendix 12.1. The variables were selected with the purpose of building the best models to predict missing values of the childbirth self-efficacy measures. I selected variables that I a priori believed to be predictive of non-response. As only respondents without any missing values can be used as predictive variables, I examined the number of missing values for each variable. All predictor variables had five missing values or below, and I handled the missing values by assigning them the most frequent response category.

Sensitivity analyses

As supplementary analyses I performed the same sensitivity analyses as described in section 5.4:

1. The multinomial logistic regression analyses were performed excluding participants who used concomitant birth and parent preparation.
2. In addition, the following per protocol analyses were performed to take adherence into account: 1) participants in the intervention group who adhered to the intervention versus all participants in the control group; 2) participants in the intervention group who adhered to the intervention versus participants in the control group who participated in both antenatal lectures. Adherence to the intervention was defined as attending all three sessions before birth and using the website.

Analyses were performed using SAS v. 9.3, SAS Institute Inc.

SUMMARY OF RESULTS

Main analyses

Results from the analyses of the effect of the intervention on childbirth self-efficacy, measured with three single items, indicated positive effects of the NEWBORN programme (table 5.5).

Table 5.5. Adjusted odds ratios (OR) (95% confidence interval (CI)) for low and neither/nor vs. high childbirth self-efficacy when using the control group as reference#.

			Low versus high self-efficacy*	Neither/nor versus high self-efficacy*
	Intervention	Control	Adjusted OR (95% CI)	Adjusted OR (95% CI)
Confidence in ability to cope at home during labour				
High self-efficacy % (n)	78.4% (519)	75.8% (513)		
Neither/nor % (n)	17.5% (116)	16.3% (110)		
Low self-efficacy % (n)	4.1% (27)	8.0% (54)		
			0.48 (0.32-0.73)	1.04 (0.81-1.33)
Confidence in own ability to make the delivery a positive experience				
High self-efficacy % (n)	93.9% (620)	91.7% (619)		
Neither/nor % (n)	5.8% (38)	7.9% (53)		
Low self-efficacy % (n)	0.3% (2)	0.4% (3)		
			0.67 (0.14-3.16)	0.72 (0.50-1.05)
Confidence in own ability to handle the birth process no matter how it turns out				
High self-efficacy % (n)	68.8% (455)	67.7% (458)		
Neither/nor % (n)	26.2% (173)	25.0% (169)		
Low self-efficacy % (n)	5.0% (33)	7.4% (50)		
			0.66 (0.44-0.98)	1.03 (0.83-1.28)

Analyses are performed using inverse probability weighting to account for missing data.
* Adjusted for parity and vulnerability.

Among women in the intervention group, 4.1% felt low confidence in their ability to cope at home during labour compared with 8.0% in the control group, i.e. they had low self-efficacy concerning staying at home in early labour. When examining differences between groups in multinomial logistic regression models using weighting to adjust for missing data, the adjusted odds ratio for low self-efficacy was 0.48 (95% CI: 0.32-0.73) in the intervention group compared with the control group, meaning that the odds for having low childbirth self-efficacy versus high childbirth self-efficacy in the intervention group was 0.48 (95% CI: 0.32-0.73) times the odds for having low childbirth self-efficacy versus high childbirth self-efficacy in the control group.

In total, only five women felt low confidence in their own ability to make the delivery a positive experience. The adjusted odds ratio for having low self-efficacy in relation to making the delivery a positive experience

was 0.67 (95% CI: 0.14-3.16) in the intervention group compared with the control group; i.e. no statistically significant difference between the intervention and control group.

Fewer women in the intervention group (5.0%) felt low confidence in own ability to handle the birth process compared with the control group (7.4%), i.e. fewer of the women receiving the NEWBORN program had low self-efficacy in relation to handling the birth process. The adjusted odds ratio for low self-efficacy was 0.66 (95% CI: 0.44-0.98) in the intervention group compared with the control group.

In addition to the analyses using inverse probability weighting to account for missing data, I performed the same analyses using complete case data. The odds ratio estimates from complete case analyses were comparable to the results in table 5.5 (results not shown).

Sensitivity analyses

Restricting the sample to women not participating in concomitant birth and parent preparation did not change the estimates regarding the effect of the intervention on the three measures of childbirth self-efficacy notably. Also, results from per-protocol analyses taking adherence to the intervention into account were generally similar to the results from the weighted analysis using the full sample (results not shown).

5.6. UPDATE OF THE SYSTEMATIC REVIEW

After conducting the NEWBORN trial, we carried out an update of the systematic review by updating the literature search and including the results from the NEWBORN trial presented in section 5.4. The literature search was updated in January 2016 and we identified 1,133 new records. Hereof, three new trials examining the effect of antenatal education in small classes versus other kinds of education (46, 78, 79) were eligible for inclusion according to the inclusion criteria presented in section 5.1. Additional results were reported (80, 81) from one of the trials already included in the review; the trial by Rouhe et al. (44).

One of the included trials examined the effect of a self-hypnosis intervention in small classes on use of epidural analgesia and obstetric outcomes (46). We used the same approach for bias assessment as we did for paper II (see description in section 5.1), and the trial was scored 'moderate risk of bias' for the objective outcomes. The dose of the intervention and control conditions in this trial was comparable to the already included trial by Werner at al. (45). I therefore conducted meta-analyses on the effect of self-hypnosis in small classes versus standard care on the reported outcomes; use of epidural analgesia, overall caesarean section, and spontaneous vaginal delivery. Pooling the results from the two trials did not alter the results reported in the published review, i.e. no significant differences between the intervention and control groups and hence no evidence of an effect of small classes versus standard care (results not shown).

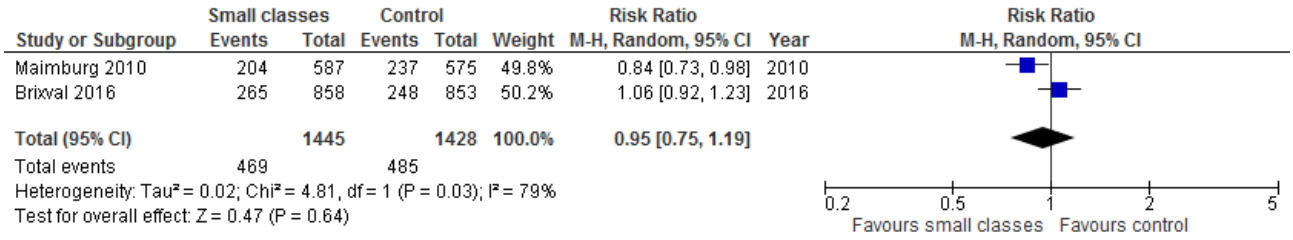
The assessment of risk of bias in the NEWBORN trial was performed by a researcher who was not part of the trial group. The NEWBORN trial was scored 'moderate risk of bias' for pain relief and obstetric outcomes (objective outcomes) and high risk of bias for childbirth self-efficacy (self-reported outcome).

META-ANALYSIS INCLUDING THE NEWBORN TRIAL

In this section I present the results from the update of the systematic review of relevance for the NEWBORN trial. As the NEWBORN trial compared an intervention comparable to the former Danish trial by Maimburg et al.; the 'Ready for Child' trial (41), we considered the trials homogeneous enough to conduct a meta-analysis of the results and thereby possibly increase the power and precision of the estimated intervention effects. To allow for some heterogeneity between trials, I used a random-effects meta-analysis.

The meta-analysis on the primary outcome of both the NEWBORN trial as well as the 'Ready for Child' trial; epidural analgesia is presented in figure 5.4. Meta-analyses on the other obstetric outcomes confirmed the results from both trials, i.e. no significant differences between the intervention and control groups (results not shown).

Figure 5.4. Meta-analysis of the effect of antenatal education in small classes versus control (standard care) on use of epidural analgesia



The results from the meta-analysis showed that after combining the results from the two trials, there was no statistically significant effect of antenatal education classes versus standard care on use of epidural analgesia (risk ratio=0.95 (95% CI: 0.75-1.19), p=0.64). The statistical heterogeneity in the meta-analysis was substantial (I^2=79%), indicating that 79% of the variability between the estimates was due to heterogeneity rather than random error.

TRIAL SEQUENTIAL ANALYSIS

Meta-analyses are at risk of producing random errors due to sparse data and multiple testing of accumulating data (82). With a relatively limited number of trials and trial participants and with an increasing number of repetitive tests, the risk of type I errors, i.e. rejection of a true null hypothesis, is large. Trial sequential analysis can be applied to assess this risk (http://www.ctu.dk/tsa/) (83).

Required information size

First, the required information size (the number of participants needed in a meta-analysis to detect or reject a pre-specified intervention effect) is estimated (84, 85). The required information size corresponds to the sample size in a randomised trial (the number of participants needed in a meta-analysis to detect or reject a pre-specified intervention effect) but the information size in a meta-analysis must also take the diversity into account, i.e. the heterogeneity between trials regarding included trial populations, interventions, and methods (D^2). The diversity can be estimated from an a priori anticipated diversity between trials or by using the diversity observed in the meta-analysis (84).

Trial sequential monitoring boundaries

On the basis of the required information size, trial sequential monitoring boundaries for benefit, harm, and futility can be constructed. If the trial sequential monitoring boundary for either benefit or harm is crossed before reaching the required information size, firm evidence for benefit or harm may be established. If the boundary is crossed, further trials may turn out to be superfluous. In contrast, if the boundary for benefit or harm is not surpassed one may conclude that it is necessary to continue with further trials before a certain intervention effect can be accepted or rejected. If the boundary for futility is crossed then it can be concluded that the anticipated intervention effects can be ruled out. It must be noted that trial sequential analysis does not in any way take into account risks of bias, and trial sequential analysis results are not reliable if high risk of bias trial results are included in the trial sequential analysis. Nevertheless, trial sequential analysis enables one to determine the statistical inference concerning cumulative meta-analysis that has not yet reached the required information size (85).

Trial sequential analysis of antenatal education in small classes on use of epidural analgesia

I conducted trial sequential analysis on the primary outcome of the NEWBORN trial and the 'Ready for Child' trial. I estimated the diversity-adjusted required information size based on an assumed use of epidural analgesia of 30% in the control group, based on the observed approximate event proportion in the control group of the largest included trial; a relative risk reduction of 20%; a type I error of 5%; a beta of 10% (power of 90%); and a D^2 of 0.79 (the observed diversity, i.e. the heterogeneity between trials observed in the trial sequential analysis) (figure 5.5).

Figure 5.5. Trial sequential analysis of antenatal education in small classes versus standard care on use of epidural analgesia

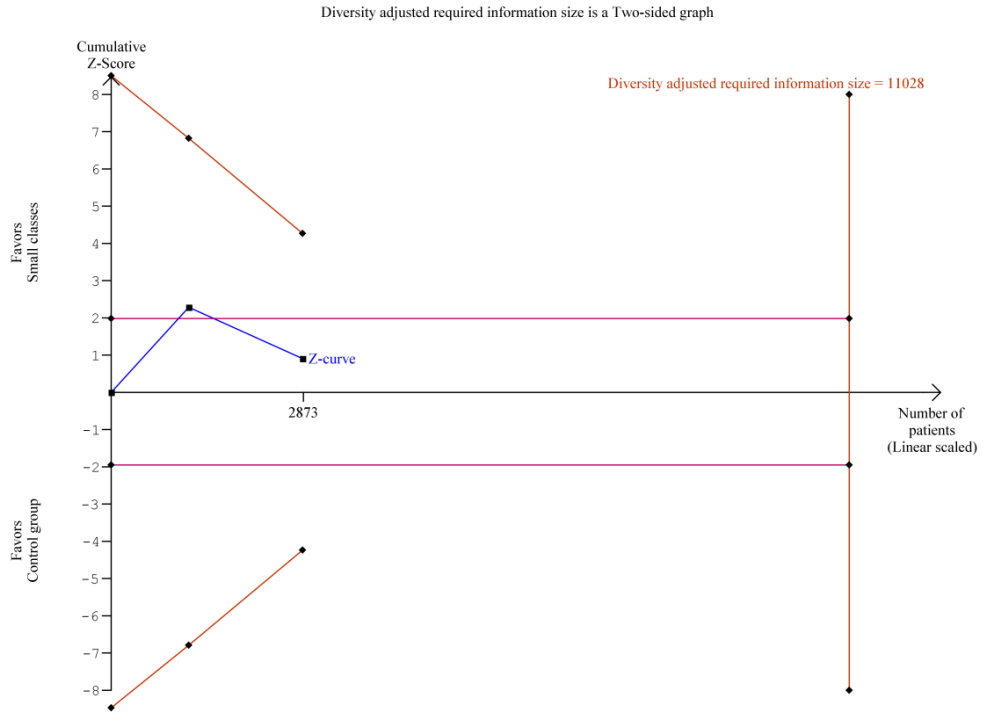


Figure 5.5 illustrates the results from the trial sequential analysis. The figure shows that 2,873 women are included in the current analysis of effect (1,162 women from the 'Ready for Child' trial and 1,711 from the NEWBORN trial). The cumulated Z-curve (blue curve) does not cross any of the trial sequential monitoring boundaries (red sloping lines), meaning that more trials are needed to establish evidence of effect. No boundaries for futility are constructed as the information fraction is too low. The diversity adjusted required information size shows that a total of 11,028 women are needed to establish evidence of an effect of antenatal education in small groups on use of epidural analgesia, i.e. a relative risk reduction of 20%.

6. DISCUSSION

The overall aim of my PhD study was to assess the effect of antenatal education in small classes on use of pain relief, obstetric interventions, and childbirth self-efficacy. I addressed this aim in the following ways: First, I conducted a systematic review examining the available evidence of the effect of antenatal education in small classes from randomised trials conducted in the Western part of the world (paper I and II). Second, I was involved in designing a large randomised trial examining the effect of antenatal education in small classes; the NEWBORN trial (paper III). Third, I validated the data source used for the obstetric outcomes (paper IV). Fourth, I examined the effect of the NEWBORN trial on use of epidural analgesia, other types of pain relief, and obstetric interventions (paper V) as well as childbirth self-efficacy (paper VI). Finally, I updated the systematic review of the effect of antenatal education in small classes and included the results from the NEWBORN trial (unpublished).

6.1. SUMMARY OF MAIN FINDINGS

THE SYSTEMATIC REVIEW

In total, 17 trials examining the effect of antenatal education in small groups in a Western setting were included in the systematic review, four of the trials assessing effects on obstetric outcomes.

There were large variations in form and content of the experimental and control conditions, trial populations, as well as outcome measures between the trials. Therefore, it was not considered appropriate to conduct meta-analyses. Based on the results from the systematic review it was not possible to draw clear conclusions on the effects of antenatal education in small classes versus other forms of antenatal education on obstetric or psycho-social outcomes. Regarding the outcomes related to use of pain relief and obstetric interventions, three of the four trials had large sample sizes (above 350 women) (41, 44, 45). Effect estimates were generally close to 1.00, indicating small or no differences between intervention and control groups. Confidence intervals were fairly narrow implying that the precision of the estimates were high and that the small effect or lack of effect did not merely reflect inadequate sample sizes or low number of events. In one trial, only 73 women were randomised (47). In this trial, effect estimates were also close to 1.00 but confidence intervals were wide due to the small sample size and low number of events (47). In conclusion, more well-conducted randomised trials with low risk of bias and adequate sample sizes are needed to establish firm evidence of the effects of antenatal education in small classes.

VALIDATION OF THE OBSTETRIC DATABASE

The completeness of the Obstetric Database, which was the source of the obstetric outcome measures used in the NEWBORN trial, was 100%. The validity of the five indicators was generally high but the indicator 'oxytocin due to dystocia' had a lower positive predictive value due to registration practices. The positive

predictive value of 'oxytocin due to dystocia' was 0.70 (95% CI: 0.57-0.80), i.e. 70% of the registrations of use of oxytocin in the Obstetric Database were also found in the medical records. The reason for the lower positive predictive value of this particular indicator was that different uses of oxytocin were registered under the same code. Due to this, I therefore chose not to use 'oxytocin due to dystocia' as an outcome measure in my analyses of effect of the NEWBORN trial.

THE NEWBORN TRIAL

In the NEWBORN trial, 1,766 women were randomised to antenatal education in small classes or the existing standard care offer consisting of two auditorium-based lectures (control group). The results from the NEWBORN trial showed no statistically significant effect of the intervention on use of epidural analgesia, other kinds of pain relief, or obstetric interventions. The odds ratio estimates were close to 1, i.e. no effect of the intervention. This indicates that the lack of intervention effect was not merely an issue of lack of power in the trial.

Analyses of the intervention effects on childbirth self-efficacy indicated positive effects of the intervention on women's confidence in their ability to cope at home during labour and confidence in own ability to handle the birth process. In the intervention group the proportion of women with low confidence in ability to cope at home during labour was half the size (4.1%) of the proportion in the control group (8.0%) corresponding to an adjusted OR of 0.48 (95% CI: 0.32-0.73). The effect of the intervention on confidence in own ability to handle the birth process was less pronounced, but clear – the adjusted OR for low self-efficacy was 0.66 (95% CI: 0.44-0.98) in the intervention group compared with the control group. There was no effect of the intervention on the women's confidence in own ability to make the delivery a positive experience (adjusted OR=0.67 (95% CI: 0.14-3.16)). The insignificant finding may reflect that only a small proportion of the women felt low confidence in own ability to make the delivery a positive experience.

UPDATE OF THE SYSTEMATIC REVIEW INCLUDING THE NEWBORN TRIAL

When updating the systematic review, one additional trial assessing the effect of antenatal education in small classes on birth-related outcomes was included, as were the results from the NEWBORN trial. The NEWBORN trial is comparable to an already included trial, and it was considered appropriate to combine results from the trials in meta-analyses and conduct trial sequential analysis on the primary trial outcome of both trials; use of epidural analgesia. The results from these meta-analyses showed no effect of antenatal education in small classes on use of pain relief or obstetric interventions. The results of the trial sequential analysis showed that more trials are needed to establish evidence of an effect (relative risk reduction of 20%) of antenatal education in small groups on use of epidural analgesia.

6.2. MAIN FINDINGS IN RELATION TO PREVIOUS STUDIES

THE SYSTEMATIC REVIEW

A systematic review by Gagnon and Sandall published in 2007 investigated the effect of structured antenatal education, including antenatal education in small classes, either to individuals or groups on a range of outcomes related to the birth process and parenthood. They concluded that the effect of general antenatal education for childbirth, parenthood, or both remained largely unknown (6). Our systematic review was focused specifically on trials investigating antenatal education in small classes (48) whereas Gagnon and Sandall used broader inclusion criteria regarding intervention and control conditions in their systematic review. For example, in addition to examining the effect of antenatal education delivered in classes, they also included educational programmes delivered to individuals (6). Hence, our review is not merely an update of the review by Gagnon and Sandall. Similar to the results from the former review, we found limited evidence from which to draw conclusions regarding the effect of antenatal education in small classes.

VALIDATION OF THE OBSTETRIC DATABASE

A previous systematic review of perinatal validation studies showed that indicators related to type of birth, e.g. caesarean section and vacuum extraction, were well reported with high sensitivities and positive predictive values. Contrarily, induction and augmentation of labour had higher degrees of underreporting (86). We found that the Obstetric Database overall had high validity concerning type of birth and epidural analgesia, but that the indicator 'oxytocin due to dystocia' had lower validity because the code used for registration covers different uses of oxytocin. The results from our validation study are thus in accordance with former validation studies in the obstetric field.

THE NEWBORN TRIAL

Only five randomised trials have examined the effect of attending antenatal education in small groups compared with other forms of education on outcomes like use of pain relief or obstetric interventions (41, 44-47) in a Western setting. One of the trials reported obstetric interventions, but the intervention itself consisted of an extra breastfeeding class (47). Hence, this trial is not considered comparable to the NEWBORN trial. Two of the trials were performed among women screened positive for fear of childbirth (44, 45) limiting generalisation of results to the general population, as most women do not suffer from fear of childbirth. Two trials assessed the effect of a self-hypnosis program, meaning that these trials were not comparable to the intervention content in the NEWBORN trial (45, 46).

One former Danish trial conducted by Maimburg and colleagues (2010); the 'Ready for Child' trial, examined the effect of antenatal education classes versus no education among primiparous women recruited among a diverse population group not limited to a high-risk population (41). Maimburg and colleagues reported a statistically significant reduced use of epidural analgesia in the intervention group,

but not of other types of pain relief and obstetric interventions. There are several possible explanations for the difference in the results between the two trials: 1) the content of the intervention programme in the 'Ready for Child' trial and the NEWBORN programme may differ. 2) We included primiparous as well as multiparous women, whereas only primiparous women were included in the 'Ready for Child' trial. It may be that the effect of antenatal education in small classes is larger among primiparous women. 3) We used 25 voluntary midwives with varying teaching experience whereas in the 'Ready for Child' trial, classes were taught by four selected midwives. Also, the midwives in the 'Ready for Child' trial may have gained more teaching experience during the trial period compared to the midwives in the NEWBORN trial as some of the midwives in our trial taught only a few classes. More experienced educators may be more confident when teaching and hence better able to include the principles of adult learning, e.g. facilitating group discussions and including the participants in the learning process. 4) The control group in the NEWBORN trial was offered auditorium-based lectures whereas no education was offered in the control group of the 'Ready for Child' trial. Although 45% of the participants in the control group in the 'Ready for Child' trial attended other forms of education this proportion was almost as high (38%) for participants in the control group of the NEWBORN trial. Hence, I consider the exposure contrast smaller in the NEWBORN trial compared with the 'Ready for Child' trial. These differences between the two trials may well have an impact on the difference in results regarding use of epidural analgesia. The exclusion of multiparous women, the use of more experienced educators, and a higher exposure contrast in the 'Ready for Child' trial may explain the finding of no effect of the NEWBORN programme compared to the effect observed in the 'Ready for Child' trial.

We found no comparable trials conducted in Western countries assessing the effect of antenatal education in small classes on the intermediate outcome in the NEWBORN trial: childbirth self-efficacy.

6.3. Discussion of mechanisms behind the findings from the NEWBORN trial

The results from the NEWBORN trial indicate that although the intervention had no effect on the use of epidural analgesia, the programme may have the potential to enhance the women's childbirth self-efficacy. As presented in figure 2.1 in section 2.4, the potential influence of antenatal education in small classes on use of epidural analgesia work through various pathways. The results point to the fact that the programme may have succeeded in affecting the first pathway, i.e. the effect of small classes on childbirth self-efficacy. It is plausible that the intervention was not comprehensive enough to impact the rest of the chain of mechanisms between self-efficacy and use of epidural analgesia. It could also be that our hypothesis that increased childbirth self-efficacy lead to reduced use of epidural analgesia is not correct. Maybe, the enhanced childbirth self-efficacy implied that women to a higher extent felt able to make their own decisions about the kind of pain relief they wanted to use in labour. It would have been an advantage if we

had been able to measure the other suggested pathways, e.g. if the woman was in active labour at the time of arrival to the labour ward as well as the women's experience of pain and anxiety. This way we could have further investigated the impact of the intervention on the intermediate pathways. Although we considered measuring these factors in the design phase it was deemed to be impossible due to practical and ethical concerns at Hvidovre Hospital. There was only a small difference in the proportion that used epidural analgesia between the intervention and control group. This may indicate that hospital practices are key factors in the decision to use pain relief. A recent Iranian trial examined the effect of an antenatal education programme focusing on increasing childbirth self-efficacy (43). They found, as we did in the NEWBORN trial, that the programme increased childbirth self-efficacy but showed no statistically significant intervention effect on delivery by caesarean section or vacuum extraction. They did not examine the effect of the programme on use of pain relief. In this trial, childbirth self-efficacy was measured by 17 questions concerning expected delivery self-efficacy (43).

6.4. METHODOLOGICAL CONSIDERATIONS

THE SYSTEMATIC REVIEW INCLUDING THE UPDATE

The systematic search of research literature in terms of a systematic review is one of the strengths of this thesis. The systematic review was carried out using the Cochrane Handbook for Systematic Reviews of Interventions as a guide (3), and the protocol of the review (50) was published before the initiation of the review itself facilitating transparency and reducing impact of author's bias (3). Extensive searches in relevant databases were performed. Trials were selected for inclusion and data extraction and evaluation of the bias risk were conducted by two independent review authors. One of the limitations of the systematic review is the great heterogeneity between trials making comparison of trial effects across trials difficult. Therefore no meta-analyses were performed in the published review. A general limitation in all systematic reviews is the risk of publication bias, implying that trials showing statistically significant findings are more likely to be published than trials reporting non-significant results (3). This might imply that there are unpublished results that might have contributed with further information on the effect of antenatal education in small classes.

The purpose of the systematic review was to guide decision-making in Western countries, and we therefore excluded trials conducted in non-Western countries. This can be seen both as a strength and as a limitation of the review. Restricting the setting implies that results cannot be generalised to a non-Western setting. On the contrary, if we had included non-Western countries the conclusions of the review might not be applicable in either setting because preparation for birth and parenthood is very sensitive to culture and contextual factors.

In future systematic reviews, it might be relevant also to look into the effect of antenatal education in small classes in a non-Western setting, and hence include trials from all countries. In the case that it is appropriate to conduct meta-analyses after new trials have been conducted and published it would be possible to supplement the overall analyses with subgroup analyses stratifying by cultural entities as sensitivity analyses.

We chose to focus primarily on evaluating evidence about the form of antenatal education, i.e. education in small classes and not the content as such. To look into the effect of content, it would be relevant to conduct a systematic review evaluating this aspect.

In the update of the systematic review, I chose to conduct meta-analyses combining the results from the trial by Maimburg and colleagues (41) and the NEWBORN trial. In spite of the differences between the trials as has been discussed in section 6.2, the trials were considered sufficiently homogeneous for conducting a meta-analysis. The statistical heterogeneity in the meta-analysis was high (I^2=79%). In the protocol for the systematic review, we defined an I^2-value of >50% as substantial (50). A high I^2-value means that the variation in the effect estimates between the two trials to a large extent is due to heterogeneity rather than random error. Substantial heterogeneity reduces the confidence of the recommendations about the intervention (87). To allow for some heterogeneity in trial characteristics in the meta-analysis, I used a random-effect model to estimate the pooled effect estimate. The random-effect method allows for some heterogeneity in the treatment effect between the trials and produces wider confidence intervals if heterogeneity is present, meaning that statistically significant effect sizes are more difficult to obtain and the random-effect model hence lead to more conservative effect estimates compared to the fixed-effect model (3).

The application of trial sequential analysis in the updated systematic review is another strength of my study because conventional meta-analyses may produce random errors due to sparse data and repetitive testing of accumulating data (82). The estimation of the required information size is dependent on the selected parameters. I chose to use the approximate observed event proportion in the control group of the largest included trial: the NEWBORN trial, for the calculation of the required information size. Also, I chose a relative risk reduction of 20% as this was the risk reduction deemed relevant for clinical practice. Choice of levels for type I and type II errors were based on conventional choices and diversity was defined as the observed diversity. Use of other parameters would lead to different required information sizes. For this reason, the choice of parameters should preferably be published in a review protocol before conducting the trial sequential analysis. Although not published, I chose the parameters prior to conducting the analysis, limiting the risk of bias.

VALIDATION OF THE OBSTETRIC DATABASE

Validation of register-based data is necessary to ensure the quality of data, and I therefore validated the information in the data source used for the obstetric outcomes; the Obstetric Database. The methodological strengths of this validation study includes performing sample size calculations, drawing a random sample of women giving birth at the hospital, and comparing the information in the register with information in the primary data source; medical records. In addition, we were two assessors; a medical student and I, who independently extracted data from the medical records to reduce the risk of recording error in the process of manual inspection of the medical records.

The results from the validation study showed that information about 'oxytocin due to dystocia' consisted of different types of use of oxytocin due to registration practices. Administration of oxytocin due to dystocia as well as oxytocin used as part of induction of labour are both registered under the same code. As presented in figure 2.1 in section 2.4, I consider use of oxytocin due to dystocia as a possible step on the pathway between antenatal education and use of epidural analgesia. Women arriving on the labour ward before being in active labour are at greater risk of being diagnosed with dystocia due to increased monitoring of labour. This mechanism was considered susceptible to be influenced by the intervention. However, use of oxytocin used as part of inductions of labour, e.g. because of prolonged pregnancy, was not considered to be modifiable by the NEWBORN programme. The discovery that the two types of usage were registered under the same code implied that I did not use administration of oxytocin as an outcome measure in the NEWBORN trial. It is likely that this discovery of the dual use of the code for oxytocin due to dystocia would not have been made if we had not carried out this validation study, and hence I consider it a strength of the thesis to have conducted and included this validation.

The Obstetric Database is a local hospital-based register. The results of the validation study might not be generalisable to other clinical databases at other hospitals as registration practices might vary between hospitals. However, the registration guidelines for the obstetric coding apply throughout the entire country which suggests that the results may be generalisable to other local clinical databases (68).

THE NEWBORN TRIAL

Design

We aimed at conducting a trial with low risk of bias in order to help guide decision-making around antenatal education. The setting at Hvidovre Hospital at the time the NEWBORN trial was initiated was ideal for investigation of this issue as the antenatal education was offered as auditorium-based lectures, although the region at the time was planning to implement birth and parent preparation in small classes.

The NEWBORN trial is the largest randomised trial to date assessing the effect of antenatal education in small classes versus auditorium-based lectures. Strengths of the trial include that a trial protocol, including

sample size calculations and a statistical analysis plan, was published (65). The publication of a trial protocol promotes proper trial implementation, reduce avoidable protocol amendments, and facilitate proper reporting and external review of the trial (88). Analyses were conducted according to the analysis plan. Other strengths of the trial design include using a computer-generated allocation sequence and performing centralised computer-based randomisation to reduce the risk of bias related to the randomisation process.

Intervention and control group

Development of the programme

The intervention programme in the NEWBORN trial was developed using a systematic framework for health promotion programme planners warranting effective decision making at each step in the intervention planning, implementation, and evaluation (67). We ensured that the form and content of the intervention programme fulfilled the Danish Health Authority's and the Capital Region's Birth Planning Committee's recommendation regarding form, i.e. antenatal education delivered in small classes and the content, i.e. antenatal education with focus on aspects related to the transition to parenthood as well as information about the delivery (7). Further, the intervention was delivered using a detailed teaching manual developed for the trial (89). This makes it possible to implement the NEWBORN programme in clinical practice and to replicate our results in future trials.

We aimed at developing a programme that would be possible to implement in an everyday clinical practice setting if proven effective. We therefore consulted politicians and service providers to determine the practical and economic level of service which they would be willing to support, hence endeavouring sustainability of the intervention after finalisation of the research project. The timeframe for the classes in the intervention group was therefore developed balancing time needed to cover included subjects adequately, and what service providers deemed a sustainable service. It is possible that provision of a more comprehensive programme would lead to larger intervention effects. However, in the current setting with limited healthcare resources, it is questionable if a more costly programme would be implemented in a real-life setting even if proven effective.

We intended that all midwives, who signed up for teaching, should attend a one-day workshop prior to the start of the intervention. However, due to practical reasons, it was not feasible to offer this to all midwives, and midwives who signed up for teaching after the onset of the intervention received a less intensive preparation. Midwives who attended the one-day workshop were likely better prepared for the facilitating role compared to midwives who merely observed a session. This may have been reflected in their ability to include the principles of adult learning, e.g. facilitating group discussions and including the participants in the learning process. This might imply that the effect of the NEWBORN programme was larger among

participants attending the classes facilitated by midwives, who had attended the one-day workshop. It would be relevant to investigate this issue in greater detail.

Choice of control group

We focused on conducting a trial using the existing standard care offer as control condition. Some trials examine the effect of different interventions without using standard care as reference. This approach implies that the results of the trial are difficult to use for decisions on change in provision of care. Using standard care as the control group is advantageous as the effect of the intervention is compared to the existing offer making decisions on change in provision of care more straightforward.

Implementation

Intention-to-treat analyses are considered preferable when examining effects of trials due to the low risk of selection bias. However, process evaluations including assessment of the implementation of the programme can contribute with further knowledge about the impact of the programme, and make it possible to distinguish between lack of effects of an intervention due to poor quality of the intervention or inadequate implementation of the intervention. Implementation fidelity is the degree to which a programme is implemented as intended by the developers and is defined by five elements: adherence to the intervention, dose (delivered and received), quality of delivery, participant responsiveness, and programme differentiation (90). To give an indication of the implementation of the programme, I used two simple measure of adherence in this thesis: 1) adherence to the intervention, defined as participating in all three sessions before birth and using the web-page and 2) adherence to the programme content in session 2, measured by asking the participants whether they had heard about the main topics of that session. For example the participants were asked whether they had heard about 'what to do at home in the early phase of labour' or 'pain relief' during the session. However, these crude measures may not provide sufficient information about whether the intervention was implemented as intended. In the NEWBORN trial, we have also measured other aspects of implementation such as *dose delivered* by asking the midwives whether they have covered the planned topics of the session. Also, we have measured *participant responsiveness* by asking the participants whether they experienced the topics as relevant.

Thorough measurement of implementation may be used to differentiate proper and improper implementation. It is outside the scope of this thesis to examine the implementation of the programme in detail, but it would be beneficial to try to understand the implementation aspects of the intervention further, i.e. to understand whether lack of success of the intervention can be partly explained by improper implementation.

Outcome measures

Epidural analgesia

We chose use of epidural analgesia as the primary trial outcome. We considered use of epidural as a proxy for coping with the delivery as is presented in our programme theory (figure 5.2, section 5.2). It would, therefore, have been relevant to additionally measure the women's ability to cope with the delivery by use of self-reported measures at the labour ward. However, due to time constraints at the labour ward this was not feasible. This information might also be collected after the delivery. However, such retrospectively collected data on coping could be influenced by information bias, i.e. that the woman's answer also depended on her actual delivery experience and performance.

Data on epidural analgesia and other obstetric outcomes were collected from a local hospital-based register: the Obstetric Database, which is used for administrative purposes and forms the basis for research conducted at the Department of Obstetrics and Gynaecology at Hvidovre Hospital. In general, register-based data is considered less prone to recall bias and missing data, and hence have higher validity than other data sources (91). The validity of Obstetric Database was confirmed by our validation study showing a positive predictive value of epidural analgesia of 100% (68).

Childbirth self-efficacy

The three items measuring the intermediate trial outcome: childbirth self-efficacy, were developed specifically for the NEWBORN trial and have not been validated. A comprehensive scale for measuring childbirth self-efficacy has been developed: the Childbirth Self-Efficacy Inventory (CBSEI) (92). CBSEI builds on Bandura's theory of self-efficacy (35) and consists of 62 items, measuring both outcome expectancy as well as efficacy expectations, and covering two stages of birth; active labour and second stage labour.

Although measuring childbirth self-efficacy by this scale might have contributed with more thorough details about the women's self-efficacy, it was not feasible to include this long scale in the questionnaires. In the NEWBORN trial, we assessed three secondary and several explorative outcome measures by use of questionnaires. Due to limited space in the questionnaires, it was prioritised to measure self-efficacy by three single items. One dimension of self-efficacy is outcome expectations, which concerns the individual's belief that a given behaviour will lead to certain outcomes. An example of this could be the women's belief that staying at home longer would decrease the risk of obstetric interventions. However, we aimed at increasing the other dimension of childbirth self-efficacy; efficacy expectations, i.e. the individual's belief in own ability to perform a specific behaviour. We therefore chose to focus on measuring efficacy expectations. As our aim was to measure the effect of our programme, we asked the women about their confidence in their ability to cope with situations that were touched upon in the session about the delivery;

importance of staying at home until the labour was in progress, the couple's possibility to influence the birth process, and what to do if the delivery did not turn out as expected (89).

In spite of the possible limitations regarding precision and consistency related to using non-validated single item questions (93), these measures may be good indicators of the women's childbirth self-efficacy because they capture the essential elements in Bandura's self-efficacy theory; confidence in own ability to perform specific behaviours.

Blinding

Blinding is a key factor for reducing bias in randomised trials. In trials with subjective outcomes effect estimates are often exaggerated when there is inadequate or unclear blinding (94). In the NEWBORN trial, it was not possible to blind participants and educators and it was therefore chosen to use an objectively measured outcome as the primary outcome, which reduces the risk of bias caused by lack of blinding (94, 95). In the communication to the personnel at Hvidovre Hospital, and in other settings, the NEWBORN trial was referred to as a general programme aiming to increase resources for birth and parenthood among expectant parents, and it was not made explicit that the primary outcome of the trial was use of pain relief. Although it is possible that the women in this trial informed the personnel at the labour ward about their intervention status in the trial, it seems unlikely that the decision to provide pain relief or perform obstetric interventions rely on this information as such decisions are made by the midwives and physicians at the labour ward and is expected to be unrelated to the intervention status. To further ensure blinding, I conducted the statistical analyses of the obstetric outcomes while blinded to intervention group. Blinding was maintained until the Steering Committee had drawn conclusions on trial effects on the primary and secondary outcomes.

Childbirth self-efficacy was self-reported and hence it was not possible to assess this outcome blinded. This may have created biased estimates because the participants were aware of their intervention status and this awareness in itself may have influenced their answers to the questionnaire. When I performed the statistical analyses of childbirth self-efficacy I was not blinded. However, the strategy for analyses was planned before the analyses were performed.

Attrition and handling of missing outcome data

The attrition for use of epidural analgesia in the NEWBORN trial was low (3%) and distributed evenly between the intervention and control group. The majority of the missing data was caused by deliveries carried out at hospitals not incorporated in the Obstetric Database system, e.g. due to moving to another city. In spite of the low attrition, I examined the possible impact of missing outcome data by conducting a 'worst case' and a 'best case' scenario analysis. In this sensitivity analysis an uncertainty interval is calculated for the intervention effect including all uncertainty due to missing data (96). Best-worst case analysis is considered an applicable sensitivity analysis when a binary outcome is missing in only a small

proportion of participants (72, 73). However, it is also argued that imputing all missing values to good or bad is too strong an assumption (97), and that results from this type of sensitivity analysis is too extreme to be realistic and should not be given much weight in the interpretation of results (98).

Attrition for childbirth self-efficacy was 24% and the proportion was fairly equal between groups. The higher attrition for the self-efficacy measure was expected as this measure was self-reported by the women. I accounted for the potential attrition bias by using inverse probability weighting in the analyses. The weighted estimates were not considerably different from complete case data indicating that attrition bias was not a substantial problem based on the assumption that data is missing at random.

Another method for handling of missing data is multiple imputation (97). Multiple imputation is usually considered more efficient in handling missing data than inverse probability weighting (99). However, inverse probability weighting is preferable in situations where all variables are missing, e.g. due to loss to follow-up, and where there are only few missing values of the predictive variables (99). I therefore considered it appropriate to use inverse probability weighting to account for missing data on childbirth self-efficacy. The inverse probability weighting method requires full report on the variables used for prediction of missing values. As there were only very few missing values I decided to assign the most frequent response category to the missing values to fill out the missing data. Another approach, I could have taken to fill out missing values was multiple imputation technique. However, considering the small amount of missing data on the predictor variables I decided that it was a sufficient approach to assign the most frequent value before conducting the inverse probability weighting analysis.

Generalisation

The NEWBORN trial was conducted at one Danish hospital which may limit the generalisability of results to other hospitals. However, the trial population was recruited among a diverse population group and not limited to a high-risk population. This increases the likelihood of results be generalisable to a general population.

A total of 19.6% of the women invited for participation in the NEWBORN trial accepted and were randomised. Although we aimed to recruit a diverse population group to the trial, the participants were predominantly primiparous women and women with a higher education level compared to the general population of Copenhagen women in the same age group (100). The high proportion of women with a university education (76%) in the trial population may imply that the women included in the trial find the auditorium-based teaching form more appealing than the general population. This could mean that the effect of the NEWBORN programme would be larger among people with a lower educational level. Moreover, the proportion receiving pain relief and obstetric interventions (except elective caesarean section) were higher among the trial population than among the total population of women giving birth at Hvidovre Hospital in 2014 (101). These discrepancies between the trial population and background

population characteristics may limit the generalisability of the trial results, and the intervention might have different effects among specific population groups, e.g. among women with a lower educational level (102).

The vast majority of the women enrolled in the NEWBORN trial had high confidence in their ability to cope with different aspects of the delivery. It is possible that there could be a ceiling effect as to the effect of the NEWBORN programme improving self-efficacy among women already highly confident in own ability to cope. Provision of the programme to women less confident might reach different conclusions.

7. CONCLUSIONS

Antenatal education is widely used and hence represents considerable costs to the healthcare system. In a healthcare system with limited resources it is important that decisions on provision of care are made on a sound basis. Preferably, decisions should be based on evidence from systematic reviews and randomised trials.

Results from our systematic review, including the update, were inconclusive as to the effect of antenatal education in small classes versus other forms of antenatal education. The trials included were too heterogeneous to conduct meta-analyses making comparisons of results between trials difficult. More well-conducted, comparable randomised trials with low risk of bias are needed to establish firm evidence of the effects of antenatal education in small classes.

The validation of the Obstetric Database showed that the database was overall complete and most of the indicators had high validity. However, the examination of validity also revealed that codes for 'oxytocin due to dystocia' covered different uses of oxytocin. This finding highlights the importance of careful consideration and evaluation of the completeness and validity of registers before using data for research.

Results from the NEWBORN trial showed no effect of an antenatal education programme delivered in small classes versus auditorium-based lectures on use of pain relief or obstetric interventions. However, although we did not succeed in reducing use of epidural analgesia among participants allocated to the NEWBORN programme, the programme seems to have had some impact on the intermediate outcome; childbirth self-efficacy.

8. IMPLICATIONS

8.1. IMPLICATIONS FOR RESEARCH

Practices on antenatal education are to a large extent based on experiences and personal beliefs by the health personnel, and there is a lack of evidence of which form of antenatal education to provide.

Based on our systematic review, it was not possible to establish evidence of an effect of antenatal education in small classes due to large variations in form and content of the experimental and control conditions, trial populations, as well as outcome measures between the trials. Hence, there is a need to conduct more high-quality, randomised trials with adequate sample sizes and transparent reporting of relevant outcome measures to establish evidence of the effect of antenatal education in small classes. Results from these trials should be included in future updates of the systematic review.

In order to replicate trials it is essential to report the content of the intervention and control conditions in detail. Future trials should focus on evaluating programmes that are likely to be implementable in an everyday clinical practice setting, if proven effective. If the purpose of a trial is to guide decision-making about provision of healthcare service, the relevant comparison group is standard care. When comparing the relative effects of two or more programmes, it is not be possible to determine the effect of the programme in relation to standard care. Future trials should, therefore, first focus on a comparison to standard care rather than comparing the relative effects of different educational programmes. However, it is important to note that standard care is different in different settings, even between hospitals in Denmark. This makes it difficult to compare trials across settings, and results of individual trials may not be generalisable to other settings.

In the systematic review I did not assess the quality of evidence for each outcome separately. GRADE (Grades of Recommendation, Assessment, Development and Evaluation) is a tool for grading the quality of evidence for each outcome separately across trials and the quality of evidence for each outcome is up-weighted or down-weighted according to the quality of the studies, directness of evidence, heterogeneity, precision of effect estimates, and risk of publication bias (3). It would be relevant in future updates of the systematic review to assess the certainty of the evidence for each outcome and thereby rate the strength of recommendations to decision-makers separately for each outcome.

As shown in this thesis, evidence of effects of antenatal education in small classes obtained from randomised trials is scarce. In this case, results obtained from non-randomised studies may be considered. One advantage of observational studies is that results may be more generalisable compared to results from randomised trials (103). The population in a trial is often more homogeneous and may not represent the general population. However, the risk of selection bias in an observational study is larger and this needs to

be considered when interpreting the results. There are methods developed to assess the effect of an intervention even in a non-randomised study, such as matching or propensity score matching, although use of these methods are not considered as effective in reducing bias as randomisation (103).

Additionally, it is possible to include non-randomised studies in systematic reviews and meta-analyses. Inclusion of non-randomised studies in systematic reviews may be needed in situations where it is necessary to provide evidence synthesis to guide policy makers and service providers in the absence of randomised trials (104). Assessment of risk of bias should also be performed in non-randomised studies, e.g. by use of the ROBINS (Risk Of Bias In Non-randomized Studies) tool (105). When combining results from randomised trials and non-randomised studies, GRADE can be used to rate the quality of evidence for each outcome separately (3, 104). Using the GRADE approach will often imply that results from non-randomised studies are given less weight in the systematic review as the risk of bias is often considered high in these studies (3).

In the NEWBORN trial, we found that only 19.6% of the invited women accepted to participate in the trial, and these were not representative of the background population regarding parity and educational level. We did not examine reasons for non-participation in detail but some of the women did not have the time to participate in an extended birth and parent preparation offer. Others did not want to participate because of the amount of time that they would need to spend on answering questionnaires. The intervention was designed to meet the recommendations from the Capital Region's Birth Planning Committee and the Danish Health Authority. In the planning of the trial, we were additionally inspired by former studies on attitudes towards antenatal education and statements from midwives from the planning group saying that many future parents request antenatal education in small classes. The value of the results of a randomised trial is larger when based on a broad population group. It would be beneficial to conduct qualitative studies to perform thorough needs assessments among the relevant target groups to assess whether the intended programme is perceived as relevant by all population groups, e.g. examine the needs among women with lower educational levels. In addition to knowledge on whether a certain intervention has an effect, policy makers and service providers require knowledge about whether the intervention is feasible and acceptable to the general population (106). This further highlights the importance of conducting thorough process evaluations.

Not only design of the intervention programme as such, but also design of the trial is important when attempting to reach specific groups. Trial researchers should consider designing a trial applicable to the population groups of interest. If the intention is to reach a broad population group it is essential to design the trial so that it appeals to various groups of people. For example, it might be that special considerations should be made to include persons with lower educational levels in future trials. Examples of things to consider are design of the invitation letter and the length and level of difficulty of questionnaires. It would be beneficial to focus on development of short, applicable and valid measures of psycho-social outcomes,

e.g. self-efficacy and stress, as long and complicated questionnaires is likely a barrier for reaching diverse trial populations. In addition to heighten participation among people with lower educational level, shorter and less complex questionnaires could potentially result in higher response and reduce attrition among all groups of participants (107, 108).

We found only a small difference in the proportion that used epidural analgesia between the intervention and control group. It might be that hospital practices play a larger role in the decision to use pain relief than individual factors. In the NEWBORN trial we focused on reaching the expecting parents and developing a program aiming to influence determinants for e.g. use of pain relief, stress, and parenting alliance, that were deemed amenable to change. If the purpose is to reduce use of epidural analgesia in the Danish birth sites it may be beneficial to conduct a randomised trial examining the effect of an intervention directed towards the health personnel.

Future research from the NEWBORN trial

In this thesis, I have reported results from the NEWBORN trial on the outcomes related to birth. The aims of antenatal education are broad and, in addition to birth-related outcomes, include outcomes related to psycho-social factors in the transition to parenthood, e.g. stress and postnatal depression. Outcomes like these are important to take into consideration when examining the effect of antenatal education. In the NEWBORN trial, we have defined the following outcomes as secondary outcomes: perceived stress, parenting stress, and parenting alliance. The results of the effect of the NEWBORN programme on these outcomes will be reported in future publications.

I found that the NEWBORN programme did not have statistically significant effects on use of epidural analgesia, other types of pain relief, or obstetric interventions in intention-to-treat analyses. My results further indicated that the results were not influenced by adherence to the intervention, defined as participation in all three sessions before birth and using the web-page. However, thorough process evaluation of interventions, including measuring different elements of implementation fidelity, is able to contribute with knowledge on the effect of the programme if implemented as intended. It is therefore important that effect evaluations of trials are accompanied by process evaluations and that these are reported. Further examination of the effect of the NEWBORN trial on different outcomes will take more thorough implementation measures into consideration, e.g. regarding the participants' perceived relevance of the topics. Further, it will be relevant to conduct analyses taking the midwife's level of preparation into account, i.e. by stratifying analyses for type of preparation (one-day workshop or less).

Also, the programme might have differential effects among certain subgroups, e.g. women with a lower educational level. It would be beneficial to conduct research focusing on the effect of the programme among these subgroups. Conducting long-term follow-up studies is also relevant. For example, it would be possible

to link participant data to Danish registers in order to investigate outcomes such as use of health care services and divorce/family break-ups.

8.2. IMPLICATIONS FOR PRACTICE

The recommendation from the Danish Health Authority is that antenatal education is delivered in small classes. These recommendations are primarily based on learning theory and studies of what future parents request and not on firm evidence from randomised trials and systematic reviews.

Based on this thesis it is not possible to guide decision-making for changes in the provision of antenatal education offer. The results from the systematic review of effect of antenatal education in small classes versus other types of education were inconclusive due to lack of comparable studies. Results from the NEWBORN trial lend some support to the recommendation of implementing antenatal education in small classes regarding increasing childbirth self-efficacy but we found no statistically significant effect of the intervention on use of pain relief or obstetric interventions. The finding that the intervention increased childbirth self-efficacy is based on explorative analyses of an intermediate trial outcome and decisions on provision of care should not be based on these results.

The NEWBORN trial showed no statistically significant effect of the intervention on the obstetric outcomes. Recommendations on whether to implement the programme in clinical practice also depend on results of the intervention effect on the secondary outcomes; perceived stress, parenting stress, and parenting alliance. Likewise, results of the effect of the programme among subgroups and results from analyses taking implementation into consideration may contribute to guide decision-making. These issues need to be investigated before recommendation of implementation of the programme in clinical practice can be validly expressed.

Randomised trials and systematic reviews are the preferable sources of evidence on which to base decisions on provision of care on. However, in the present situation where insufficient evidence exists as to which form of antenatal education to provide, non-randomised observational studies may contribute with knowledge on the provision of care. The risks of bias are larger in non-randomised studies and results must be interpreted in light of this limitation.

9. ENGLISH SUMMARY

Background

In health care systems with limited resources, these should be used to provide the forms of health care proven to be most effective. The decisions about what care to provide should be based on results from the most reliable sources of evidence, i.e. systematic reviews with meta-analyses and randomised trials.

Antenatal education aims to help prospective parents prepare for childbirth and parenthood. Although antenatal education is widely used and hence represents considerable costs to the healthcare system, the effect of the form and the content of the education are poorly evaluated. Shifts in practice appear to be based on tendencies, professional beliefs, political wishes, and economic considerations rather than on solid evidence from systematic reviews and trials favouring one form over the other. In many Danish birth sites, the practice has been to provide antenatal education as auditorium-based lectures, although the Danish Health Authority recommends that antenatal education is delivered in small classes. According to learning theorists, people learn more effectively in a group setting, where they for example have the opportunity to observe others' perspectives, to interact regularly, and to supplement one another. Also, education in small classes provides opportunity for group activities in which the participants become actively engaged which may lead to a better learning outcome. Hence, antenatal education in small classes may be beneficial over auditorium-based lectures. However, evidence of an effect of antenatal education in small classes is sparse.

Objectives

The overall aim of this thesis was to assess the effect of antenatal education in small classes on use of pain relief, obstetric interventions, and childbirth self-efficacy. This aim was translated into the following objectives: 1) to assess the current evidence for the effect of antenatal education in small classes versus other types of education in a systematic review (paper I and II); 2) to design a randomised trial examining the effects of antenatal education in small classes on use of epidural analgesia, other types of pain relief, obstetric interventions, and psycho-social outcomes – the NEWBORN trial (paper III); 3) to test the validity of the data source used for the obstetric outcomes; the Obstetric Database (paper IV); 4) to examine the effect of the NEWBORN trial on use of epidural analgesia, other types of pain relief, and obstetric interventions (paper V); 5) to examine the effect of the NEWBORN trial on the intermediate trial outcome; childbirth self-efficacy (paper VI); and 6) to update the evidence for the effect of antenatal education in small classes on obstetric outcomes allowing new trials to be included in the systematic review.

Methods

I conducted a systematic review examining the effect of antenatal education in small classes versus other types of education. The review was conducted in accordance with the methods recommended by the

Cochrane Collaboration, including publishing a review protocol before embarking the review itself and scoring of risk of bias. I included randomised trials conducted in the Western world, irrespective of language, publication year, publication type, and publication status.

I examined the completeness and validity of the Obstetric Database by linking data from all women registered in the Obstetric Database as having given birth in 2013 to the National Patient Register. Validity of five selected indicators from the Obstetric Database was assessed using medical records as a golden standard. I calculated proportion of agreement, sensitivity, specificity, and positive and negative predictive values for each indicator. We were two assessors who independently reviewed medical records and I calculated the inter-rater reliability as the proportion of agreement and Cohen's Kappa coefficient.

In the NEWBORN trial, 1,766 women, planned to give birth at Hvidovre Hospital in 2013-2014, were randomised to the intervention group (antenatal education in small groups three times during pregnancy and one time after the delivery) or control group consisting of standard care (two auditorium-based lectures during pregnancy). The primary outcome of the trial was use of epidural analgesia. Intervention effects were examined using logistic regression models (pain relief and obstetric interventions) and multinomial logistic regression models (childbirth self-efficacy) with adjustment for stratification variables. Methods used to adjust for missing data were best-worst case analysis (epidural analgesia) and inverse probability weighting (childbirth self-efficacy).

I updated the systematic review allowing inclusion of new trials and results from the NEWBORN trial. I conducted meta-analyses where appropriate and conducted trial sequential analysis to estimate the required information size needed to establish firm evidence for effect of antenatal education in small classes on use of epidural analgesia.

Results

In total, 17 trials examining the effect of antenatal education in small groups were included in the systematic review, hereof four trials assessing effects on obstetric outcomes. Due to heterogeneity in intervention- and control groups, meta-analyses were not performed and conclusions on provision of antenatal education cannot be drawn based on this systematic review.

The completeness of the Obstetric Database was 100%. The validity of the five indicators was generally high but the indicator 'oxytocin due to dystocia' had a lower positive predictive value due to registration practices implying that different uses of oxytocin is registered under the same code.

Results from the NEWBORN trial showed no effect of the intervention on use of epidural analgesia, other kinds of pain relief, or obstetric interventions. Analyses of the intervention effect on childbirth self-efficacy indicated positive effects of the intervention on women's confidence in their ability to cope at home during labour and confidence in own ability to handle the birth process.

The results from the update of the systematic review showed no effect of antenatal education in small classes on use of pain relief or obstetric interventions. The results of the trial sequential analysis showed that more trials are needed to establish evidence of an effect of antenatal education in small groups on use of epidural analgesia.

Conclusions

Results from the systematic review were inconclusive as to the effect of antenatal education in small classes versus other forms of antenatal education. More well-conducted randomised trials with low risk of bias are needed to establish firm evidence of the effects of antenatal education in small classes. Results from the NEWBORN trial indicated that the antenatal education programme in small classes increased childbirth self-efficacy but showed no statistically significant effect on use of pain relief or obstetric interventions. The results of this thesis led to no clear recommendations on which form of antenatal education should be provided in clinical practice.

10. DANSK RESUMÉ

Baggrund

I et sundhedsvæsen med begrænsede ressourcer, herunder det danske, bør de tilgængelige ressourcer anvendes på de mest effektive behandlinger. Beslutninger om, hvilke behandlinger, der skal tilbydes, bør derfor baseres på resultater fra de mest pålidelige kilder til evidens: systematiske reviews med meta-analyser og randomiserede forsøg.

Formålet med fødsels- og forældreforberedelse er at hjælpe kommende forældre med at forberede sig til fødslen og forældrerollen. Fødselsforberedelse er et udbredt tilbud og repræsenterer betydelige omkostninger for sundhedsvæsenet, men området er præget af mangelfuld evaluering af effekt af fødselsforberedelsens form og indhold. Ændringer i den anvendte praksis i Danmark synes at være baseret på tendenser, faglige overbevisninger, politiske ønsker og økonomiske overvejelser fremfor evidens baseret på systematiske reviews og forsøg. På mange danske fødesteder har praksis været at tilbyde fødselsforberedelse i form af auditorium-baserede forelæsninger, selv om Sundhedsstyrelsen anbefaler, at fødselsforberedelse foregår i mindre grupper. Ifølge læringsteoretikere lærer folk mere effektivt, når de indgår i en mindre gruppe, hvor de eksempelvis har mulighed for at høre andres perspektiver, interagere med hinanden og supplere hinanden. Undervisning i mindre grupper giver mulighed for gruppeaktiviteter, hvor deltagerne bliver aktivt involveret, og dette kan medføre bedre indlæring. Det er derfor sandsynligt, at fødselsforberedelse i mindre hold kan være mere fordelagtigt end auditorium-baserede forelæsninger. Dog er dokumentationen for en effekt af fødselsforberedelse i mindre hold mangelfuld.

Formål

Det overordnede formål med mit ph.d.-studie var at vurdere effekten af fødselsforberedelse i mindre hold på brug af smertelindring, obstetriske indgreb og fødsels self-efficacy. Dette overordnede formål blev oversat til følgende delformål: 1) at vurdere den aktuelle evidens for effekten af fødselsforberedelse i mindre hold kontra andre former for fødselsforberedelse i et systematisk review (artikel I og II); 2) at designe et randomiseret forsøg med det formål at undersøge effekterne af fødselsforberedelse i mindre hold på brug af epiduralblokade, andre former for smertelindring, obstetriske indgreb, og psykosociale faktorer – projekt Nyfødt (artikel III); 3) at undersøge validiteten af datakilden brugt til de obstetriske effektmål; Obstetrisk Database (artikel IV); 4) at undersøge effekten af Nyfødt på brug af epiduralblokade, andre former for smertelindring og obstetriske indgreb (artikel V); 5) at undersøge effekten af Nyfødt på det intermediære effektmål: fødsels self-efficacy (artikel VI) og 6) at opdatere evidensen for effekten af fødselsforberedelse i mindre hold på obstetriske effektmål ved at inkludere nye forsøg i det systematiske review.

Metoder

Jeg udarbejdede et systematisk review af effekten af fødselsforberedelse i mindre hold kontra andre former for fødselsforberedelse. Reviewet blev gennemført i overensstemmelse med de metoder, der er anbefalet af Cochranesamarbejdet, herunder publicering af en protokol før igangsættelse af selve reviewet samt vurdering af risiko for bias i de inkluderede forsøg. Jeg inkluderede randomiserede forsøg, der var gennemført i den vestlige verden, uanset sprog, udgivelsesår, publikationstype og hvorvidt forsøget var publiceret eller ej.

Jeg undersøgte kompletheden og validiteten af Obstetrisk Database ved at udtrække data fra alle kvinder, der i Obstetrisk Database er registreret til at have født i 2013 og koble disse til Landspatientregisteret. Validiteten af fem udvalgte indikatorer fra Obstetrisk Database blev vurderet ved at bruge de elektroniske fødejournaler som reference. Jeg beregnede andel af overensstemmelse, sensitivitet, specificitet samt positive og negative prædiktive værdier for hver indikator. Vi var to personer, der uafhængigt af hinanden gennemgik fødejournalerne og registrerede indikatorerne, og jeg beregnede inter-rater reliabiliteten som andelen af overensstemmelse mellem registreringerne og ved Cohens Kappa koefficient.

I projekt Nyfødt blev 1.766 kvinder, der var planlagt til at skulle føde på Hvidovre Hospital i perioden 2013-2014, randomiseret til interventionsgruppen (fødselsforberedelse i små hold tre gange i graviditeten og én gang efter fødslen) eller kontrolgruppen bestående af standardtilbuddet på Hvidovre Hospital (to auditorium-baserede forelæsninger under graviditeten). Forsøgets primære effektmål var brug af epiduralblokade – andre former for smertelindring og obstetriske interventioner blev behandlet som eksplorative effektmål. Fødsels self-efficacy blev behandlet som et intermediært effektmål. Effekten af interventionen blev undersøgt ved hjælp af logistiske regressionsmodeller (smertelindring og obstetriske indgreb) og multinomiale logistiske regressionsmodeller (fødsels self-efficacy) med justering for stratificeringsvariablene: sårbarhed og paritet. Metoder brugt til at justere for manglende data var 'best-worst case analyse (epiduralblokade) og inverse probability weighting (fødsels self-efficacy).

Jeg opdaterede det systematiske review og inkluderede resultater fra nye forsøg samt resultater fra projekt Nyfødt. Jeg gennemførte meta-analyser og udførte trial sequential analysis for at estimere antallet af randomiserede kvinder, der kræves for at kunne etablere klar evidens for effekten af fødselsforberedelse i mindre hold på brug af epiduralblokade.

Resultater

I alt blev 17 forsøg, der undersøgte effekten af fødselsforberedelse i mindre hold, inkluderet i det systematiske review. Af disse var der fire forsøg, der undersøgte effekter af interventioner på obstetriske effektmål. På grund af stor variation i indholdet i interventions- og kontrolgrupper, blev der ikke udført meta-analyser, og der kan ikke drages konklusioner om, hvilken form for fødselsforberedelse, der er mest effektiv på baggrund af det systematiske review.

Kompletheden af Obstetrisk Database var 100%. Validiteten af de fem indikatorer var generelt høj; dog havde 'oxytocin på grund af dystoci' en lavere positiv prædiktiv værdi end de øvrige indikatorer. Registreringspraksis indebærer, at forskellige anvendelser af oxytocin registreres under samme kode.

Resultater fra projekt Nyfødt viste ingen statistisk signifikant effekt af interventionen på brug af epiduralblokade, andre former for smertelindring eller obstetriske indgreb. Analyserne af interventionens effekt på fødsels self-efficacy indikerede positive effekter af interventionen på kvindernes tro på egen evne til at klare sig hjemme i den tidlige del af fødslen, samt på deres tro på egen evne til at håndtere fødslen uanset, hvordan den udvikler sig.

Resultaterne fra opdateringen af det systematiske review viste ingen effekt af fødselsforberedelse i mindre hold på brug af smertelindring eller obstetriske indgreb. Resultaterne af trial sequential analysis viste, at der er behov for flere forsøg for at etablere dokumentation for en effekt af fødselsforberedelse i små grupper på brugen af epiduralblokade.

Konklusion

Det var ikke muligt at konkludere noget entydigt om effekten af fødselsforberedelse i mindre hold ud fra det systematiske review, og der er behov for flere velgennemførte randomiserede forsøg med lav risiko for bias for at etablere evidens for effekten af fødselsforberedelse i mindre hold. Resultater fra *projekt Nyfødt* viste, at fødselsforberedelsesprogrammet i mindre hold øgede fødsels self-efficacy, men forsøget viste ingen statistisk signifikante effekter på brug af smertelindring eller obstetriske indgreb. Resultaterne af denne afhandling førte ikke til nogen klare anbefalinger om, hvilken form for fødselsforberedelse, der bør implementeres i klinisk praksis.

11. REFERENCES

1. Cochrane A. Effectiveness and Efficiency: Random Reflections on Health Services London: The Nuffield Provicial Hospitals Trust; 1972.
2. Guyatt G, Cairns J, Churchill D, et al. Evidence-based medicine: A new approach to teaching the practice of medicine. JAMA : the journal of the American Medical Association. 1992;268(17):2420-5.
3. Higgins Julian P.T. SG. Cochrane Handbook for Systematic Reviews of Interventions. West Sussex, England: The Cochrane Collaboration and John Wiley & Sons Ltd.; 2011.
4. Shea BJ, Hamel C, Wells GA, Bouter LM, Kristjansson E, Grimshaw J, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. Journal of clinical epidemiology. 2009;62(10):1013-20.
5. Odgaard-Jensen J, Vist GE, Timmer A, Kunz R, Akl EA, Schunemann H, et al. Randomisation to protect against selection bias in healthcare trials. Cochrane database of systematic reviews (Online). 2011(4):Mr000012.
6. Gagnon AJ, Sandall J. Individual or group antenatal education for childbirth or parenthood, or both. Cochrane database of systematic reviews (Online). 2007(3):CD002869.
7. Danish Health Authority. Recommendations for antenatal care (*in Danish*). Copenhagen: 2009.
8. Danish Health Authority. Recommendations for antenatal care (*in Danish*). Copenhagen: 2013.
9. National Board of Health. Status for the regions' and municipalities' offers for pregnant and women in labour with focus on outpatient deliveries and early discharges *(in Danish).* 2011 [Accessed June 29, 2016]. Available from: http://www.sum.dk/Aktuelt/Nyheder/Sundhedspolitik/2011/April/~/media/Filer%20-%20dokumenter/Svangreomsorg/Sundhedsstyrelsens_redegoerelse_svangreomsorg.ashx.
10. Nutbeam D. Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century. Health Prom Int. 2000;15(3):259-67.
11. Knowles MS, Holton III EF, Swanson RA. The adult learner: The definitive classic in adult education and human resource development: Routledge; 2014.
12. Illeris K. Contemporary theories of learning: learning theorists... in their own words: Routledge; 2009.
13. Nolan ML. Information giving and education in pregnancy: a review of qualitative studies. The Journal of perinatal education. 2009;18(4):21-30.
14. Deliveries and complications *(in Danish)*. 2013 [accessed March 11, 2016]. Available from: http://www.esundhed.dk/sundhedsregistre/MFR/Sider/MFR03.aspx.
15. Anim-Somuah M, Smyth RM, Jones L. Epidural versus non-epidural or no analgesia in labour. Cochrane database of systematic reviews (Online). 2011(12):CD000331.
16. Roberts CL, Tracy S, Peat B. Rates for obstetric intervention among private and public patients in Australia: population based descriptive study. BMJ (Clinical research ed). 2000;321(7254):137-41.

17. Howell CJ, Kidd C, Roberts W, Upton P, Lucking L, Jones PW, et al. A randomised controlled trial of epidural compared with non-epidural analgesia in labour. BJOG : an international journal of obstetrics and gynaecology. 2001;108(1):27-33.
18. Jones L, Othman M, Dowswell T, Alfirevic Z, Gates S, Newburn M, et al. Pain management for women in labour: an overview of systematic reviews. Cochrane database of systematic reviews (Online). 2012;3:CD009234.
19. Poignant M, Hjelmstedt A, Ekeus C. Indications for operative delivery between 1999-2010 and induction of labor and epidural analgesia on the risk of operative delivery--a population based Swedish register study. Sexual & reproductive healthcare : official journal of the Swedish Association of Midwives. 2012;3(4):129-34.
20. Greenwell EA, Wyshak G, Ringer SA, Johnson LC, Rivkin MJ, Lieberman E. Intrapartum temperature elevation, epidural use, and adverse outcome in term infants. Pediatrics. 2012;129(2):e447-54.
21. Nguyen US, Rothman KJ, Demissie S, Jackson DJ, Lang JM, Ecker JL. Epidural analgesia and risks of cesarean and operative vaginal deliveries in nulliparous and multiparous women. Maternal and child health journal. 2010;14(5):705-12.
22. Zimmer EZ, Jakobi P, Itskovitz-Eldor J, Weizman B, Solt I, Glik A, et al. Adverse effects of epidural analgesia in labor. European journal of obstetrics, gynecology, and reproductive biology. 2000;89(2):153-7.
23. O'Mahony F, Hofmeyr GJ, Menon V. Choice of instruments for assisted vaginal delivery. Cochrane database of systematic reviews (Online). 2010(11):CD005455.
24. Hansen AK, Wisborg K, Uldbjerg N, Henriksen TB. Risk of respiratory morbidity in term infants delivered by elective caesarean section: cohort study. BMJ (Clinical research ed). 2008;336(7635):85-7.
25. Tracy SK, Tracy MB. Costing the cascade: estimating the cost of increased obstetric intervention in childbirth using population data. BJOG : an international journal of obstetrics and gynaecology. 2003;110(8):717-24.
26. Lederman RP, Lederman E, Work BA, Jr., McCann DS. The relationship of maternal anxiety, plasma catecholamines, and plasma cortisol to progress in labor. American journal of obstetrics and gynecology. 1978;132(5):495-500.
27. Alehagen S, Wijma K, Wijma B. Fear during labor. Acta obstetricia et gynecologica Scandinavica. 2001;80(4):315-20.
28. Carlsson IM, Hallberg LR, Odberg Pettersson K. Swedish women's experiences of seeking care and being admitted during the latent phase of labour: a grounded theory study. Midwifery. 2009;25(2):172-80.

29. Holmes P, Oppenheimer LW, Wen SW. The relationship between cervical dilatation at initial presentation in labour and subsequent intervention. BJOG : an international journal of obstetrics and gynaecology. 2001;108(11):1120-4.
30. Rahnama P, Ziaei S, Faghihzadeh S. Impact of early admission in labor on method of delivery. International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics. 2006;92(3):217-20.
31. Jackson DJ, Lang JM, Ecker J, Swartz WH, Heeren T. Impact of collaborative management and early admission in labor on method of delivery. Journal of obstetric, gynecologic, and neonatal nursing : JOGNN / NAACOG. 2003;32(2):147-57; discussion 58-60.
32. Bailit JL, Dierker L, Blanchard MH, Mercer BM. Outcomes of women presenting in active versus latent phase of spontaneous labor. Obstetrics and gynecology. 2005;105(1):77-9.
33. Mold JW, Stein HF. The cascade effect in the clinical care of patients. The New England journal of medicine. 1986;314(8):512-4.
34. Bandura A. Self-efficacy mechanism in human agency. Am Psychol. 1982;37(2):122-47.
35. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev. 1977;84(2):191-215.
36. Lowe NK. Explaining the pain of active labor: the importance of maternal confidence. Research in nursing & health. 1989;12(4):237-45.
37. Lowe NK. Individual variation in childbirth pain. Journal of Psychosomatic Obstetrics & Gynecology. 1987;7(3):183-92.
38. Salomonsson B, Gullberg MT, Alehagen S, Wijma K. Self-efficacy beliefs and fear of childbirth in nulliparous women. Journal of psychosomatic obstetrics and gynaecology. 2013;34(3):116-21.
39. Lowe NK. Self-efficacy for labor and childbirth fears in nulliparous pregnant women. Journal of psychosomatic obstetrics and gynaecology. 2000;21(4):219-24.
40. Carlsson IM, Ziegert K, Nissen E. The relationship between childbirth self-efficacy and aspects of well-being, birth interventions and birth outcomes. Midwifery. 2015.
41. Maimburg RD, Vaeth M, Durr J, Hvidman L, Olsen J. Randomised trial of structured antenatal training sessions to improve the birth process. BJOG : an international journal of obstetrics and gynaecology. 2010;117(8):921-8.
42. Beebe KR, Lee KA, Carrieri-Kohlman V, Humphreys J. The effects of childbirth self-efficacy and anxiety during pregnancy on prehospitalization labor. Journal of obstetric, gynecologic, and neonatal nursing : JOGNN / NAACOG. 2007;36(5):410-8.
43. Taheri Z, Mazaheri MA, Khorsandi M, Hassanzadeh A, Amiri M. Effect of Educational Intervention on Self-efficacy for Choosing Delivery Method among Pregnant Women in 2013. Int J Prev Med. 2014;5(10):1247-54.

44. Rouhe H, Salmela-Aro K, Toivanen R, Tokola M, Halmesmaki E, Saisto T. Obstetric outcome after intervention for severe fear of childbirth in nulliparous women - randomised trial. BJOG : an international journal of obstetrics and gynaecology. 2013;120(1):75-84.
45. Werner A, Uldbjerg N, Zachariae R, Rosen G, Nohr EA. Self-hypnosis for coping with labour pain: a randomised controlled trial. BJOG : an international journal of obstetrics and gynaecology. 2013;120(3):346-53.
46. Downe S, Finlayson K, Melvin C, Spiby H, Ali S, Diggle P, et al. Self-hypnosis for intrapartum pain management in pregnant nulliparous women: a randomised controlled trial of clinical effectiveness. BJOG : an international journal of obstetrics and gynaecology. 2015;122(9):1226-34.
47. Duffy EP, Percival P, Kershaw E. Positive effects of an antenatal group teaching session on postnatal nipple pain, nipple trauma and breast feeding rates. Midwifery. 1997;13(4):189-96.
48. Brixval CS, Axelsen SF, Lauemøller SG, Andersen SK, Due P, Koushede V. The effect of antenatal education in small classes on obstetric and psycho-social outcomes-a systematic review. Systematic reviews. 2015;4(1):1-9.
49. Brixval CS KV, Axelsen SF, Andersen SK, Due P. The effect of antenatal education classes on obstetric and psycho-social outcomes – a systematic review: PROSPERO; 2013. Available from: http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42013004319 [accessed March 29, 2016].
50. Brixval CS, Axelsen SF, Andersen SK, Due P, Koushede V. The effect of antenatal education in small classes on obstetric and psycho-social outcomes: a systematic review and meta-analysis protocol. Systematic reviews. 2014;3:12.
51. Members and partners of OECD. Available from: http://www.oecd.org/about/membersandpartners/ [accessed March 28, 2016].
52. Feinberg ME, Kan ML. Establishing family foundations: intervention effects on coparenting, parent/infant well-being, and parent-child relations. Journal of family psychology : JFP : journal of the Division of Family Psychology of the American Psychological Association (Division 43). 2008;22(2):253-63.
53. Kistin N, Benton D, Rao S, Sullivan M. Breast-feeding rates among black urban low-income women: effect of prenatal education. Pediatrics. 1990;86(5):741-6.
54. Ickovics JR, Kershaw TS, Westdahl C, Magriples U, Massey Z, Reynolds H, et al. Group prenatal care and perinatal outcomes: a randomized controlled trial. Obstetrics and gynecology. 2007;110(2 Pt 1):330-9.
55. Le HN, Perry DF, Stuart EA. Randomized controlled trial of a preventive intervention for perinatal depression in high-risk Latinas. Journal of consulting and clinical psychology. 2011;79(2):135-41.
56. Schulz MS, Cowan CP, Cowan PA. Promoting healthy beginnings: a randomized controlled trial of a preventive intervention to preserve marital quality during the transition to parenthood. Journal of consulting and clinical psychology. 2006;74(1):20-31.

57. Westney OE, Cole OJ, Munford TL. The effects of prenatal education intervention on unwed prospective adolescent fathers. Journal of adolescent health care : official publication of the Society for Adolescent Medicine. 1988;9(3):214-8.
58. Forster D, McLachlan H, Lumley J, Beanland C, Waldenstrom U, Amir L. Two mid-pregnancy interventions to increase the initiation and duration of breastfeeding: a randomized controlled trial. Birth (Berkeley, Calif). 2004;31(3):176-82.
59. Maycock B, Binns CW, Dhaliwal S, Tohotoa J, Hauck Y, Burns S, et al. Education and support for fathers improves breastfeeding rates: a randomized controlled trial. Journal of human lactation : official journal of International Lactation Consultant Association. 2013;29(4):484-90.
60. Rossiter JC. The effect of a culture-specific education program to promote breastfeeding among Vietnamese women in Sydney. International journal of nursing studies. 1994;31(4):369-79.
61. Brugha TS, Wheatley S, Taub NA, Culverwell A, Friedman T, Kirwan P, et al. Pragmatic randomized trial of antenatal intervention to prevent post-natal depression by reducing psychosocial risk factors. Psychological medicine. 2000;30(6):1273-81.
62. Noel-Weiss J, Rupp A, Cragg B, Bassett V, Woodend AK. Randomized controlled trial to determine effects of prenatal breastfeeding workshop on maternal breastfeeding self-efficacy and breastfeeding duration. Journal of obstetric, gynecologic, and neonatal nursing : JOGNN / NAACOG. 2006;35(5):616-24.
63. Lara MA, Navarro C, Navarrete L. Outcome results of a psycho-educational intervention in pregnancy to prevent PPD: a randomized control trial. Journal of affective disorders. 2010;122(1-2):109-17.
64. Ortiz Collado MA, Saez M, Favrod J, Hatem M. Antenatal psychosomatic programming to reduce postpartum depression risk and improve childbirth outcomes: a randomized controlled trial in Spain and France. BMC pregnancy and childbirth. 2014;14:22.
65. Koushede V, Brixval CS, Axelsen SF, Lindschou J, Winkel P, Maimburg RD, et al. Group-based antenatal birth and parent preparation for improving birth outcomes and parenting resources: Study protocol for a randomised trial. Sexual & reproductive healthcare : official journal of the Swedish Association of Midwives. 2013;4(3):121-6.
66. Koushede V. Project Newborn – Preparation for Birth and Parenthood. 2014. Available from: http://www.interventionsforskning.dk/userfiles/files/nyfdt/projektprotokol-version-23_2.pdf. [cited 2015 March 11th].
67. Bartholomew LK, Parcel GS, Kok G, Gottlieb NH, Fernandez ME. Planning health promotion programs: an intervention mapping approach: John Wiley & Sons; 2011.
68. Brixval CS, Thygesen LC, Johansen NR, Rorbye C, Weber T, Due P, et al. Validity of a hospital-based obstetric register using medical records as reference. Clin Epidemiol. 2015;7:509-15.

69. ClinicalTrials.gov. NewPreBP: Project Newborn -Preparation for Birth and Parenthood 2012 [updated February 2015; accessed March 17, 2015]. Available from: https://clinicaltrials.gov/ct2/results?term=NCT01672437.
70. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977;33(1):159-74.
71. Little RJA. A Test of Missing Completely at Random for Multivariate Data with Missing Values. Journal of the American Statistical Association. 1988;83(404):1198-202.
72. White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. BMJ (Clinical research ed). 2011;342:d40.
73. Hollis S. A graphical sensitivity analysis for clinical trials with non-ignorable missing binary outcome. Stat Med. 2002;21(24):3823-34.
74. Brixval CS, Axelsen SF, Thygesen LC, Due P, Koushede V. Antenatal education in small classes may increase childbirth self-efficacy: Results from a Danish randomised trial. Sexual & Reproductive Healthcare.
75. Kwak C, Clayton-Matthews A. Multinomial logistic regression. Nursing research. 2002;51(6):404-10.
76. Seaman SR, White IR. Review of inverse probability weighting for dealing with missing data. Stat Methods Med Res. 2013;22(3):278-95.
77. Rothman KJ, Greenland, S., Lash, L.L. Modern Epidemiology. 2nd ed. Philadelphia, USA: Lippincott Williams and Wilkins 2008.
78. Bittner A, Peukert J, Zimmermann C, Junge-Hoffmeister J, Parker LS, Stobel-Richter Y, et al. Early intervention in pregnant women with elevated anxiety and depressive symptoms: efficacy of a cognitive-behavioral group program. J Perinat Neonatal Nurs. 2014;28(3):185-95.
79. Daley-McCoy C, Rogers M, Slade P. Enhancing relationship functioning during the transition to parenthood: a cluster-randomised controlled trial. Archives of women's mental health. 2015;18(5):681-92.
80. Rouhe H, Salmela-Aro K, Toivanen R, Tokola M, Halmesmaki E, Ryding EL, et al. Group psychoeducation with relaxation for severe fear of childbirth improves maternal adjustment and childbirth experience--a randomised controlled trial. Journal of psychosomatic obstetrics and gynaecology. 2015;36(1):1-9.
81. Rouhe H, Salmela-Aro K, Toivanen R, Tokola M, Halmesmaki E, Saisto T. Life satisfaction, general well-being and costs of treatment for severe fear of childbirth in nulliparous women by psychoeducative group or conventional care attendance. Acta obstetricia et gynecologica Scandinavica. 2015;94(5):527-33.
82. Brok J, Thorlund K, Gluud C, Wetterslev J. Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses. Journal of clinical epidemiology. 2008;61(8):763-9.

83. Thorlund K, Engstrøm J, Wetterslev J, Brok J, Imberger G, Gluud C. User manual for trial sequential analysis (TSA) Copenhagen Trial Unit, 2011 [Accessed March 25, 2016]. Available from: http://www.ctu.dk/tsa/files/tsa_manual.pdf.
84. Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. BMC Med Res Methodol. 2014;14:120.
85. Wetterslev J, Thorlund K, Brok J, Gluud C. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. Journal of clinical epidemiology. 2008;61(1):64-75.
86. Lain SJ, Hadfield RM, Raynes-Greenow CH, Ford JB, Mealing NM, Algert CS, et al. Quality of data in perinatal population health databases: a systematic review. Med Care. 2012;50(4):e7-20.
87. Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ : British Medical Journal. 2003;327(7414):557-60.
88. Chan AW, Tetzlaff JM, Gotzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ (Clinical research ed). 2013;346:e7586.
89. Axelsen SF, Brixval CS, Koushede V. Course material for teachers: NEWBORN - preparation for birth and parenthood. 2012.
90. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. Implementation science : IS. 2007;2:40.
91. Sorensen HT, Sabroe S, Olsen J. A framework for evaluation of secondary data sources for epidemiological research. International journal of epidemiology. 1996;25(2):435-42.
92. Lowe NK. Maternal confidence for labor: development of the Childbirth Self-Efficacy Inventory. Research in nursing & health. 1993;16(2):141-9.
93. Bowling A. Just one question: If one question works, why ask several? Journal of epidemiology and community health. 2005;59(5):342-5.
94. Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. BMJ (Clinical research ed). 2008;336(7644):601-5.
95. Savovic J, Jones HE, Altman DG, Harris RJ, Juni P, Pildal J, et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. Ann Intern Med. 2012;157(6):429-38.
96. Gamble C, Hollis S. Uncertainty method improved on best-worst case analysis in a binary meta-analysis. Journal of clinical epidemiology. 2005;58(6):579-88.
97. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ (Clinical research ed). 2009;338:b2393.
98. Unnebrink K, Windeler J. Sensitivity analysis by worst and best case assessment: Is it really sensitive? Drug Inf J. 1999;33(3):835-9.

99. Seaman SR, White IR, Copas AJ, Li L. Combining multiple imputation and inverse-probability weighting. Biometrics. 2012;68(1):129-37.
100. Statistics Denmark. Population highest completed education (*in Danish*). 2014 [accessed March 16, 2015]. Available from: http://www.statistikbanken.dk/statbank5a/default.asp?w=1422.
101. Hvidovre Hospital. Births in numbers (*in Danish*). 2014 [accessed August 7, 2015]. https://www.hvidovrehospital.dk/afdelinger-og-klinikker/foedeomraadet/om-afdelingen/Sider/Foedsler-i-tal.aspx.
102. Brixval CS, Thygesen LC, Axelsen SF, Gluud C, Winkel P, Lindschou J, et al. Effect of antenatal education in small classes versus standard auditorium-based lectures on use of pain relief during labour and of obstetric interventions: results from the randomised NEWBORN trial. BMJ Open. 2016;6(6):e010761.
103. D'Agostino RB, Kwan H. Measuring effectiveness. What to expect without a randomized control group. Med Care. 1995;33(4 Suppl):As95-105.
104. Reeves BC, Higgins JP, Ramsay C, Shea B, Tugwell P, Wells GA. An introduction to methodological issues when including non-randomised studies in systematic reviews on the effects of interventions. Research synthesis methods. 2013;4(1):1-11.
105. Sterne JAC, Higgins JPT, Reeves BC on behalf of the development group for ROBINS-I: a tool for assessing Risk Of Bias In Non-randomized Studies of Interventions, Version 7 March 2016. [accessed June 29, 2016]. Available from: https://sites.google.com/site/riskofbiastool/.
106. Moat KA, Lavis JN, Wilson MG, Rottingen JA, Barnighausen T. Twelve myths about systematic reviews for health system policymaking rebutted. Journal of health services research & policy. 2013;18(1):44-50.
107. Bergkvist L, Rossiter JR. The Predictive Validity of Multiple-Item Versus Single-Item Measures of the Same Constructs. Journal of Marketing Research. 2007;44(2):175-84.
108. Edwards PJ, Roberts I, Clarke MJ, Diguiseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires. Cochrane database of systematic reviews (Online). 2009(3):Mr000008.

12. APPENDIX

12.1. DESCRIPTION OF QUESTIONNAIRE-BASED VARIABLES USED

Variable	Description	Measurement time-point and data source	Used in paper
Allocation	Randomisation variable (intervention group versus control group).	Randomisation	V, VI
Vulnerability	Vulnerability defined by the woman's general practitioner (yes versus no). Criteria for vulnerability: former or current psychiatric disorder, current life crisis, victims of violence, adverse psycho-social background, current or former eating disorder, former suicide attempt, doubts about parenting skills, or age below 18 years.	Before baseline	V, VI
Parity	Parity defined by the woman's general practitioner (primiparous versus multiparous).	Before baseline	V, VI
Age	Woman's age at birth. Calculated as date of child's birth minus date at own birth.	Obstetric Database	V
Educational level	Woman's self-reported educational level measured by question: "What is your highest completed education?" by ticking a list with the following response categories: Not completed an education; primary school or similar; high school-level; vocational school; short tertiary education, e.g. healthcare assistant; medium tertiary education, e.g. primary school teacher, bachelor degree from the university; higher tertiary education, e.g. doctor; other, please state. I categorised the "other" category according to the other categories where possible. Otherwise, the category was set to missing. The educational level was dichotomised into ≤medium tertiary education versus higher tertiary education.	Baseline	V
Occupational status	Woman's occupational status. Self-reported by the question: "Do you have a paid job at the time?" (yes versus no).	Baseline	VI
Body Mass Index	The woman's pre-pregnancy weight and height measured by the general practitioner and converted into Body Mass Index (kg/m^2).	Obstetric Database	V
Living with child's father	Self-reported by the woman by ticking the response category "Living with the child's father" in the question: "Which grown-ups do you live with?"	Baseline	V

Planned pregnancy	Self-reported by the woman. Question: Is this pregnancy planned, partly planned or not planned" Planned pregnancy was dichotomised into: planned (yes or partly) or not planned.	Baseline	V
Self-rated physical health status	Self-reported by the woman. Question: "How would you describe your physical health status altogether?" Response categories: Excellent, very good, good, poor, very poor. Self-rated mental health was dichotomised into excellent/very good versus good, poor, very poor.	Baseline	V, VI
Self-rated mental health status	Self-reported by the woman. Question: "How would you describe your mental health status altogether?" Response categories: Excellent, very good, good, poor, very poor. Self-rated mental health was dichotomised into excellent/very good versus good, poor, very poor.	Baseline	V, VI
Feeling stressed	Self-reported by the woman. Question: "Do you feel stressed?" Response categories: "no; yes, a little; yes, moderately; yes, a lot". Stress was dichotomised into no versus yes, a little; yes, moderately; yes, a lot.	Baseline	V, VI
Antenatal/postnatal depressive symptomatology	Measured by the Edinburgh Postnatal Depression Scale (EPDS) [1] self-reported by the woman. EPDS consists of ten items. All answers are added together to a sum score with a potential range from 0-30. The cutpoint EPDS≥13 was used as an indicator of antenatal/postnatal depressive symptomatology.	Baseline, 37 weeks of gestation	V
Perceived stress	Measured by the Perceived Stress Scale (PSS) [2] self-reported by the woman. PSS consists of ten items. All answers are added together to a sum score with a potential range from 0-40.	Baseline	V
Childbirth self-efficacy	Measured by three single items: 1) I believe that I will feel confident at home once labour has begun (e.g. before going to the labour ward), 2) I believe that I can contribute to making the birth a good experience 3) I believe that I will be able to handle the birth process no matter how it turns out. All three items had the following response categories: totally agree, agree, neither/nor, disagree, and totally disagree. I trichotomised the responses in the following categories: high self-efficacy (totally agree, agree), neither/nor, and low self-efficacy (disagree, totally	37 weeks of gestation	VI

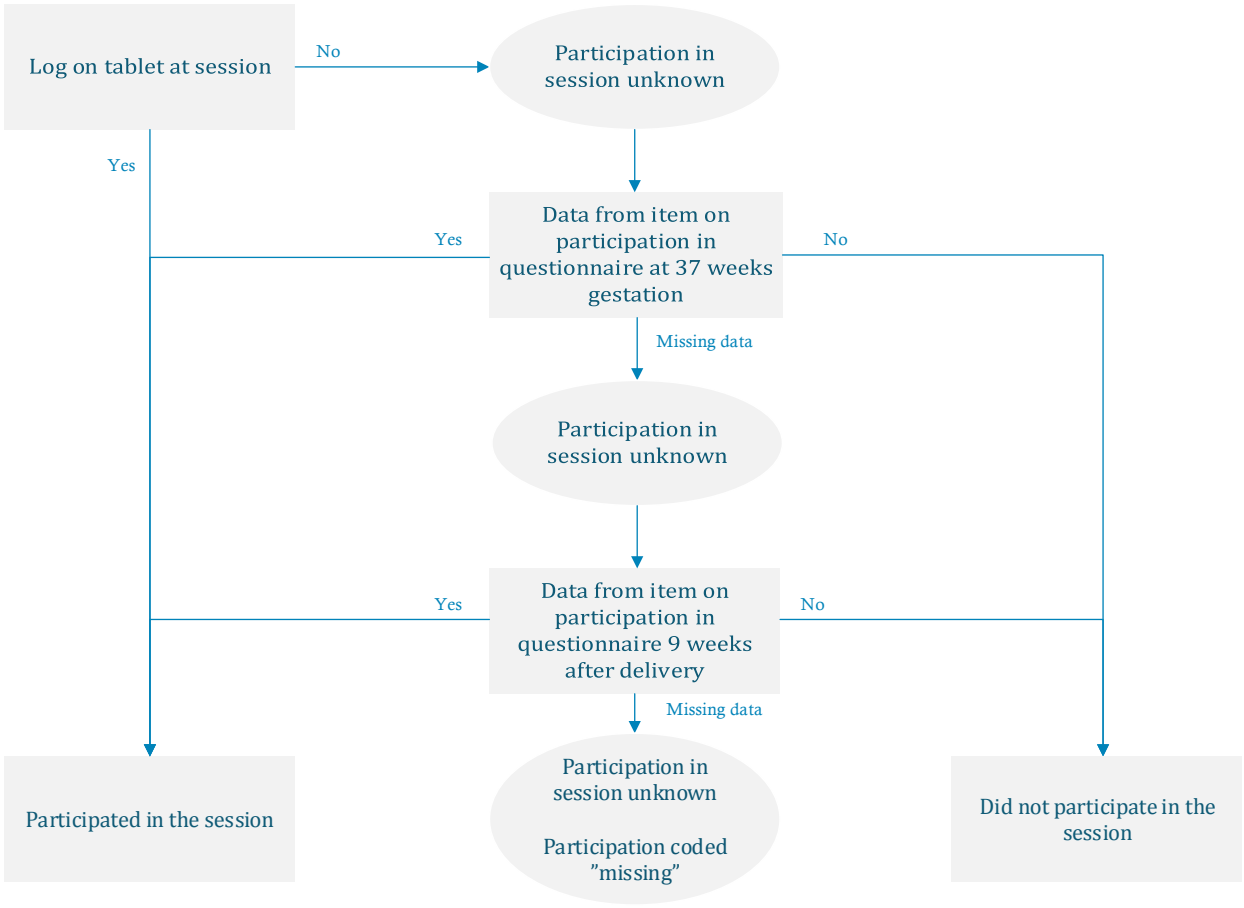
	disagree).		
Participation in antenatal education	Measured by the question; "During pregnancy: which offers of antenatal birth preparation have you participated in? (you are welcome to tick more than one box)" with the following response categories: None; lecture about breastfeeding at Hvidovre Hospital; lecture about the delivery at Hvidovre Hospital; the NEWBORN sessions; other, please state. Participation in the lectures at Hvidovre Hospital and participation in other antenatal education were coded according to the response to this question.	37 weeks of gestation	V, VI
Adherence to the intervention	Measured by a combination of data collected by tablet after each session and responses to questions from the questionnaires. Please see appendix 12.2 for a thorough description of the categorisation of adherence.	Tablet after sessions, 37 weeks of gestation, 9 weeks after delivery	V, VI
Adherence to the programme topics in session 2	Assessed by tablet-based questionnaires. Questions: "Have you heard about xxx today". The topics asked about were: 'expectations in relation to birth', 'what to do at home in the early phase of labour', 'the normal course of labour, 'pain relief and coping strategies', 'partner support during labour', and 'when there is a need to intervene in labour'. Participants could answer yes, no, or don't know.	Tablet after session 2	V

1. Cox, J.L., J.M. Holden, and R. Sagovsky, *Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale.* Br J Psychiatry, 1987. **150**: p. 782-6.
2. Cohen, S., T. Kamarck, and R. Mermelstein, *A global measure of perceived stress.* J Health Soc Behav, 1983. **24**(4): p. 385-96.

12.2. CLASSIFICATION OF ADHERENCE TO THE INTERVENTION

Step 1: Participation in separate sessions

Figure 12.1: Illustration of classification of participation in each session before delivery



I coded participation in all three sessions "yes" if all sessions were coded "yes". In cases where the women had missing data on at least one of the sessions participation was coded "missing". Participation in all three sessions were coded "no" if one or more of the sessions were coded "no".

Step 2: Use of web-page

I used data from the item: "How often do you visit the web-page 'Netværket Nyfødt'?" collected in the questionnaire at 37 weeks gestation. Use of web-page was coded "yes" if participants answered one of the following categories: several times a day; every day; 5-6 times a week; 2-4 times a week; 1-4 times a month; less than once a month. Use of web-page was coded "no" if the woman answered the category: "I have not signed up for 'Netværket Nyfødt'. Non-response to the item was coded "missing".

Step 3: Overall adherence to the intervention

I coded adherence "yes" if the woman participated in all three sessions before delivery and was coded "yes" in use of web-page. Otherwise, adherence was coded "no" or "missing".

PAPERS

Paper I

SYSTEMATIC REVIEWS

PROTOCOL Open Access

The effect of antenatal education in small classes on obstetric and psycho-social outcomes: a systematic review and meta-analysis protocol

Carina Sjöberg Brixval*, Solveig Forberg Axelsen, Stig Krøger Andersen, Pernille Due and Vibeke Koushede

Abstract

Background: The aims of antenatal education contain both outcomes related to pregnancy, birth and parenthood. Both content and methods of antenatal education have changed over time without evidence of effects on relevant outcomes. The effect of antenatal education in groups, with participation of a small number of participants, may differ from the effect of other forms of antenatal education. The latest Cochrane review, assessed as up-to-date in 2007, concluded that the effect of antenatal education for childbirth or parenthood or both remains largely unknown. This systematic review and meta-analysis aims to assess the effects of antenatal education in small groups on obstetric as well as psycho-social outcomes.

Methods/design: Eligible studies include individually randomized as well as cluster-randomized trials irrespective of language, publication year, publication type, and publication status. Only interventions carried out in the Western world will be considered in this review. We will search the databases Medline, EMBASE, CENTRAL, CINAHL, Web of Science, and PsycINFO using relevant search terms. Two independent review authors will extract data and assess risk of bias. Results will be presented as structured summaries of the included trials. A meta-analysis will be conducted. We will assess heterogeneity by using both the Chi-squared test and the I-squared statistic, and conduct subgroup analysis separately for various intervention types.

Discussion: In healthcare systems with limited resources evidence of the effectiveness of services provided is important for decision making, and there is a need for policy makers to implement changes in healthcare systems based on scientific evidence. The effectiveness of antenatal education in small classes is still questioned. Therefore an up-to-date systematic review is needed.
This systematic review protocol was registered within the International Prospective Register of Systematic Reviews (PROSPERO) as number CRD42013004319.

Keywords: Antenatal education classes, Obstetric, Labor, Birth, Parenting, Parenthood, Psycho-social, Stress, Postnatal depression

Background

Antenatal education is offered to pregnant women in most high-income countries, more recently also to expecting fathers. Antenatal education has the overall aim of providing expecting parents with strategies for dealing with pregnancy, childbirth and parenthood [1]. More specific aims include influencing health behavior, increasing confidence in women's ability to give birth, informing about pain relief, and promoting breastfeeding.

Antenatal education has been sensitive to opinions and trends, and has undergone marked changes over time. In some periods the focus has mainly been on maternal exercise and relaxation techniques, in other periods on antenatal education in small classes with group discussions, and in others again on lectures in large auditoriums with information on childbirth and breastfeeding. Likewise, the number of sessions has also changed over time due to financial and structural changes in the healthcare sector [2]. All these changes have occurred without sound

* Correspondence: cabr@niph.dk
National Institute of Public Health, University of Southern Denmark, Copenhagen C, Denmark

evidence of the effect of antenatal education on outcomes relevant to healthcare providers as well as expecting parents [3].

The current evidence points to the importance of interacting with fellow learners and the learning environment in order to obtain new competencies [4]. In antenatal education classes that have a small number of participants it may be possible to create an environment which enables expecting parents to discuss feelings and concerns. Furthermore, it may enhance their awareness of own resources and provide them with problem-solving strategies that enhance important competencies to cope with birth and parenthood [5]. However, this approach has not been subjected to thorough scrutiny.

A previous systematic review by Gagnon and Sandall [3] investigated the effect of structured antenatal education either to individuals or groups on a range of outcomes both related to the birth process and parenthood and concluded that the effect of general antenatal education for childbirth or parenthood or both remains largely unknown [3]. However, since then more randomized trials have been conducted and the results from these trials might alter this conclusion. An updated review is therefore due.

In healthcare systems with limited resources evidence of the effectiveness of services provided is important for decision making, and there is a need for policy makers to implement changes in healthcare systems based on scientific evidence [6]. An up-to-date systematic review is needed in order to raise evidence for the effectiveness of antenatal education in small classes compared to no or other forms of education. The aims of antenatal education are numerous and various and therefore the purpose of our systematic review will be to assess the effects of antenatal education in small classes on various outcomes related to obstetric as well as psycho-social factors. Therefore, the specific research question is:

In expecting parents in a Western setting: What are the effects of antenatal education in small classes on obstetric and psycho-social outcomes compared to no intervention, treatment as usual, or other types of educational programs?

Methods and design

In accordance with the guidelines, this systematic review protocol was registered within the International Prospective Register of Systematic Reviews (PROSPERO) on 11 April 2013 (registration number CRD42013004319).

Types of studies and participants

Eligible studies will include individually randomized trials and cluster-randomized trials irrespective of language, publication year, publication type, and publication status to assess the effect of antenatal education in small classes.

Preparation for birth and parenthood are very dependent on culture and contextual factors, such as the organization of the health system. Therefore we will exclude trials taking place in developing countries and only include studies conducted in Western countries. We define Western countries as OECD membership countries [7]. We will include studies of pregnant women and/or their partners that have provided their informed consent to participation in the given trial.

Types of interventions

The experimental intervention must be delivered as an antenatal educational program offered by an educator to groups consisting of more than one individual/couple, related to the birth of an infant and/or preparation for parenthood.

The control intervention can be either no intervention, treatment as usual, or other types of educational programs. If two programs are compared, the most intensive will be considered the experimental intervention.

Co-interventions are allowed but must be equally delivered in both the experimental and control arm.

Types of outcome measures

Results must include quantitative data for outcomes measured. Both outcomes assessed as self-reported, via registries, or reported by a health professional will be accepted. If outcomes are measured more than once during follow-up, we will use the measurement shortly after the intervention ends and at the longest follow-up to consider the intervention effect.

The primary outcomes are: proportion of participants who receive pain relief during labor; proportion of participants who receive obstetric interventions; mean endpoint score in scales assessing psychological and social adjustment to parenthood; and proportion of participants with symptoms of antenatal and postnatal depression and anxiety (measured as defined by the trial).

The secondary outcomes are: knowledge acquisition; maternal sense of control/active decision making during labor and birth; partner involvement at birth; breastfeeding success; infant care abilities; and social support (all measured as defined by the trial).

Search methods for identification of studies

Extensive searches will be performed by an information specialist (SKA). Medline, EMBASE, CENTRAL, CINAHL, Web of Science, and PsycINFO will be searched. The terms will include the following: antenatal, prenatal, education, parent preparation. Searches will be limited to randomized trials. Search words will be adapted to each database. An example is given in Table 1.

In addition, we will search for relevant trials in citations from identified papers and former reviews. There will be

Table 1 Medline search strategy, modified as needed for use in other databases

Search[a]	Medline
1	(antenatal OR prenatal OR pregnancy OR birth OR childbirth OR (labor OR labour) OR obstetric OR (delivery OR deliveries))
2	(education OR "parent education" OR preparation OR "parent preparation" OR "early intervention")
3	1 AND 2

[a]Filters: Refined by *randomized controlled trial*, *humans*.

no language or publication date restriction. The searches will be re-run just before the final analyses and further studies retrieved for inclusion.

Selection of studies and data extraction

We will conduct the selection of studies in two steps. First two of the three review authors (CSB, VK, SFA) will independently perform the initial screening of all titles and abstracts to determine eligibility of all studies identified through the literature search. Next two of the three review authors (CSB, VK, SFA) will independently assess the full papers identified as meeting inclusion criteria or where definite decision on exclusion could not be made from screening titles and abstracts. Any discrepancies between the two review authors will be resolved through consultation with a third review author (PD).

A PRISMA flow diagram of progress will be completed for the selection process.

Data from the included papers will be extracted to summary tables containing information on: population, study design, interventions, theoretical framework, outcomes, type of effect analysis, results, and information for assessment of the risk of bias.

Assessment of risk of bias in included studies

Two review authors (CSB, VK) will independently assess the included trials according to a predefined risk of bias scoring key [8] in order to determine the likely presence or absence of biases which might have affected the internal validity of the trials. Any discrepancies between the two review authors will be resolved through consultation with a third review author (PD).

The scoring key includes the following characteristics:

- Selection bias: randomization sequence generation and allocation concealment.
- Performance bias: assessment of blinding of participants, personnel, and outcome assessment.
- Attrition bias: assessment of systematic differences in withdrawal of study participants between the groups compared.
- Reporting bias: assessment of systematic differences between reported and unreported findings. It will be assessed whether a trial protocol exists and whether outcomes in the published trial have been reported in a pre-specified way.
- Other sources of bias: assessment of whether sample size and power calculations of the trial are based on the reported outcome.

First, each trial will be evaluated according to each of the above-mentioned bias domains as either 'low', 'unclear', or 'high risk of bias'. Second, the trials will be will be rated by an overall risk of bias. All trials rated as 'low risk of bias' in all domains will be scored 'overall low risk of bias'. All other trials will be scored 'overall high risk of bias'.

Data analysis

Structured summaries of the included trials will be presented, structured around type of intervention, intervention content, population characteristics and type of outcome. Intervention effects from the included trials will be calculated and presented as risk ratios (for dichotomous outcomes) or standardized mean differences (for continuous outcomes) with 95% confidence intervals and two-sided *P* values for each outcome.

We anticipate that there will be limited scope for meta-analysis because of the range of different outcomes reported from trials on antenatal care. However, where trials have used the same type of intervention and comparator, with the same outcome measure, we will pool the results using a random-effects meta-analysis, with standardized mean differences for continuous outcomes and risk ratios for dichotomous outcomes, and calculate 95% confidence intervals and two sided *P* values for each outcome. Outcomes measured by ordinal scales are analyzed according to the method presented in the included trial.

In studies where the effects of clustering have not been taken into account, we will adjust the standard deviations for the design effect. Heterogeneity will be assessed using both the Chi-squared test and the I-squared statistic. We will consider an I-squared value greater than 50% indicative of substantial heterogeneity. We will conduct sensitivity analyses based on study quality.

If the necessary data are available, subgroup analyses will be done separately for various intervention types: specific class content (for example, childbirth, parenting), size of classes in the intervention, number of antenatal education sessions, timing of classes, specific teaching approaches (for example, didactic, experiential), or effects in specific population groups (for example, socio-demographic factors, parity). Likewise, we will do subgroup analyses based on risk of bias; comparing effects of interventions with 'overall high risk of bias' and interventions with 'overall low risk of bias'.

Trial sequential analysis will be done for significant results [9]. This analysis reduces the risk of type I errors,

which may occur in meta-analysis due to the repeated testing of significance.

Statistical analyses will be based on intention-to-treat and calculated using the Cochrane statistical package, Review Manager (RevMan 2003).

Discussion

This systematic review will assess the literature on the effect of antenatal education in small classes on both obstetric and psycho-social outcomes and compare with no or other forms of education. Since the aims of antenatal education are various, the present review will evaluate the effect on a broad range of outcomes in order to capture any relevant effect.

In 2007 a systematic review by Gagnon and Sandall was conducted [3] evaluating the effect of both individual and group antenatal education for childbirth or parenthood. They concluded that high-quality evidence was lacking, and that the effects of antenatal education are largely unknown. However, since 2007 more randomized trials have been conducted and results from these trials might alter this conclusion. The present systematic review will partly update the results from Gagnon and Sandall's systematic review. We, however, will limit our focus to trials of antenatal education in small classes conducted in a Western setting.

Antenatal education is dependent on culture as well as organization of the healthcare system. Since the purpose of this review is to contribute to guidance of decision making in the Western world, only trials conducted in Western countries will be included in this systematic review. Comparing effects of antenatal education across very different healthcare systems may give a misleading view of the effects in a Western setting.

In many countries antenatal education have changed dramatically over time without letting evidence guide decisions for these changes. The results from this systematic review will help guide policy makers in making evidence-based decisions on the field of antenatal education.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

CSB, VK, SFA, and PD developed the design of the protocol, drafted the manuscript, and will participate in extracting data and interpreting the results. SKA has developed the search strategy and will perform the literature search. All authors have read and approved the manuscript.

Authors' information

CSB: Master of Science in Public Health, PhD student on a large randomized trial (the NEWBORN trial) evaluating the effect of a structured antenatal education program.
VK: Midwife, MPH, PhD; principal investigator of the NEWBORN trial.
SFA: Pharmaconomist, Exam.pharm.cons, MPH; research assistant on the NEWBORN trial.
SKA: Information specialist; has expertise in literature searching.
PD: Professor, dr.med.Sci.; workpackage-chair of Child Intervention Research as part of The Centre of Intervention Research and research director of the research program for Child and Adolescent Health at The National Institute of Public Health, University of Southern Denmark.

Acknowledgements

We would like to thank Jane Lindschou from the Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Denmark for her valuable input in designing this systematic review.
The NEWBORN trial is funded by the Danish Cancer Society and this systematic review is related to the NEWBORN trial. The Danish Cancer Society does not have any influence on the decision to submit this manuscript.

Received: 9 December 2013 Accepted: 6 February 2014
Published: 13 February 2014

References

1. Ahlden I, Ahlehagen S, Dahlgren LO, Josefsson A: **Parents' expectations about participating in antenatal parenthood education classes.** *J Perinat Educ* 2012, **21:**11–17.
2. Fabian HM, Radestad IJ, Waldenstrom U: **Childbirth and parenthood education classes in Sweden. Women's opinion and possible outcomes.** *Acta Obstet Gynecol Scand* 2005, **84:**436–443.
3. Gagnon AJ, Sandall J: **Individual or group antenatal education for childbirth or parenthood, or both.** *Cochrane Database Syst Rev (Online)* 2007, **3:**CD002869.
4. Illeris K (Ed): *Contemporary Theories of Learning: Learning Theorists…in their own words.* New York: Routledge; 2009.
5. Ip WY, Tang CS, Goggins WB: **An educational intervention to improve women's ability to cope with childbirth.** *J Clin Nurs* 2009, **18:**2125–2135.
6. Moat KA, Lavis JN, Wilson MG, Rottingen JA, Barnighausen T: **Twelve myths about systematic reviews for health system policymaking rebutted.** *J Health Serv Res Policy* 2013, **18:**44–50.
7. **Members and partners of OECD**. http://www.oecd.org/about/membersandpartners/.
8. Higgins Julian PT, Green S: *Cochrane Handbook for Systematic Reviews of Interventions.* Chichester: The Cochrane Collaboration and John Wiley & Sons Ltd.; 2011.
9. Wetterslev J, Thorlund K, Brok J, Gluud C: **Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis.** *J Clin Epidemiol* 2008, **61:**64–75.

doi:10.1186/2046-4053-3-12
Cite this article as: Brixval *et al.*: The effect of antenatal education in small classes on obstetric and psycho-social outcomes: a systematic review and meta-analysis protocol. *Systematic Reviews* 2014 **3**:12.

Paper II

Brixval et al. Systematic Reviews (2015) 4:20
DOI 10.1186/s13643-015-0010-x

SYSTEMATIC REVIEWS

RESEARCH Open Access

The effect of antenatal education in small classes on obstetric and psycho-social outcomes - a systematic review

Carina Sjöberg Brixval*, Solveig Forberg Axelsen, Stine Glenstrup Lauemøller, Stig Krøger Andersen, Pernille Due and Vibeke Koushede

Abstract

Background: The aims of antenatal education are broad and encompass outcomes related to pregnancy, birth, and parenthood. Both form and content of antenatal education have changed over time without evidence of effects on relevant outcomes. The effect of antenatal education in groups, with participation of a small number of participants, may differ from the effect of other forms of antenatal education due to, for example, group dynamic. The objective of this systematic review is to assess the effects of antenatal education in small groups on obstetric as well as psycho-social outcomes.

Methods: Bibliographic databases (Medline, EMBASE, CENTRAL, CINAHL, Web of Science, and PsycINFO) were searched. We included randomized and quasi-randomized trials irrespective of language, publication year, publication type, and publication status. Only trials carried out in the Western world were considered in this review. Studies were assessed for bias using the Cochrane risk of bias tool. Results are presented as structured summaries of the included trials and as forest plots.

Results: We identified 5,708 records. Of these, 17 studies met inclusion criteria. Studies varied greatly in content of the experimental and control condition. All outcomes were only reported in a single or a few trials, leading to limited or uncertain confidence in effect estimates. Given the heterogeneity in interventions and outcomes and also the high risk of bias of studies, we are unable to draw definitive conclusions as to the impact of small group antenatal education on obstetric and psycho-social outcomes.

Conclusions: Insufficient evidence exists as to whether antenatal education in small classes is effective in regard to obstetric and psycho-social outcomes. We recommend updating this review following the emergence of well-conducted randomized controlled trials with a low risk of bias.

Systematic review registration: PROSPERO CRD42013004319

Keywords: Antenatal education classes, Obstetric, Labor, Birth, Parenting, Parenthood, Psycho-social, Postnatal depression, Systematic review, Randomized trials

* Correspondence: cabr@niph.dk
National Institute of Public Health, University of Southern Denmark, Copenhagen, Denmark

Background

Antenatal education is offered to pregnant women in most high-income countries, more recently also to expecting fathers. Antenatal education has the overall aim of providing expecting parents with strategies for dealing with pregnancy, childbirth, and parenthood [1]. More specific aims include increasing knowledge, e.g., on antenatal and postnatal depression, the birth process, pain relief and obstetric interventions, promoting breast feeding, and increasing confidence in women's ability to give birth as well as becoming parents. Also, information imparted on health promotion and risk reduction is an important aim of antenatal education. Meeting others in the same situation and developing social networks is another aim of antenatal classes [2].

Antenatal education is well-established in many Western countries, but the type and arrangement of the education is debated. Antenatal education has been sensitive to opinions and trends, and both form and content have undergone marked changes over time. During certain periods, practice has been centered on antenatal education in small classes with group discussions - in others, the practice has been lectures in large auditoriums. Also, the content has varied greatly. Topics like, for example, breathing and/or relaxation techniques have been included and left out of antenatal education intermittently. Due to financial and structural changes in the health care sector, the numbers of antenatal education sessions have also changed over time [2]. All these changes have occurred without evidence of an effect of antenatal education on outcomes relevant to expecting parents as well as health care providers [3].

Current evidence points to the importance of interacting with fellow learners and the learning environment in order to obtain new competencies [4]. In antenatal education classes with a small number of participants, it is possible to create an environment which enables expecting parents to discuss feelings and concerns. Furthermore, it may enhance their awareness of their own resources and provide them with problem-solving strategies that enhance important competencies to cope with birth and parenthood [5]. However, this approach has not been subject to thorough scrutiny.

In health care systems with limited resources, policy makers should be able to make informed decisions about health care priorities based on scientific evidence [6]. According to service providers, insecure parents use the health care services beyond indication. Janicke and Finney suggested that the use of pediatric services is a function of perceived parental stress and low self-efficacy related to coping with life demands [7]. Antenatal education in small classes may increase parenting resources leading to health care cost savings in the long term although the immediate expenses are larger for small classes than for auditorium lectures.

A previous systematic review by Gagnon and Sandall from 2007 investigated the effect of structured antenatal education, including antenatal education in small classes, either to individuals or groups on a range of outcomes both related to the birth process and parenthood and concluded that the effect of general antenatal education for childbirth or parenthood or both remains largely unknown [3].

A systematic review is needed in order to assess currently available evidence for the effectiveness of antenatal education in small classes compared to no or other forms of education. The aims of antenatal education are numerous and vary in nature. Therefore, the objective of this systematic review is to assess the effectiveness of antenatal education in small classes on obstetric and psycho-social outcomes compared to standard care or other types of educational programs using randomized trials from Western countries.

Methods

We carried out this systematic review using the Cochrane Handbook for Systematic Reviews of Interventions as a guide [8]. We published our methods as a protocol before conducting the review [9] and registered the review within the International Prospective Register of Systematic Reviews (PROSPERO) (registration number CRD42013004319 http://www.crd.york.ac.uk/PROSPERO/register_new_review.asp?RecordID=4319&UserID=2668). This systematic review is reported according to the PRISMA statement [10] [see Additional file 1].

Search strategy

Extensive searches were performed by an information specialist (SKA). The databases Medline, EMBASE, CENTRAL, CINAHL, Web of Science, and PsycINFO were searched. Search words were adapted to each database. Searches were limited to randomized trials. The full search strategy for each database is provided in Additional file 2. We searched for trials in two rounds: at the beginning of the review process and just before completion. The final search was performed 5 March 2014.

We also searched for relevant trials in citations from identified papers and former reviews. In addition, unpublished results from included trials were obtained from contact with authors. There was no language or publication date restriction.

Eligibility criteria

Eligible studies included individually randomized trials, including quasi-randomized trials, and cluster-randomized trials.

Setting

Preparation for birth and parenthood are dependent on culture and contextual factors, such as the organization of the health care system. Trials taking place in developing countries have therefore been excluded and only trials conducted in the Western world - defined as OECD membership countries - are included [11].

Participants

We have included studies of pregnant women and/or their partners that have provided their informed consent to participation in the given trial or where descriptions in the papers indicate the participants' consent to randomization.

Experimental and control conditions

The experimental conditions in the included trials must be delivered as an antenatal educational program offered by an educator to groups consisting of more than one individual/couple but including less than 20 individuals, related to delivery and/or preparation for parenthood. The control conditions in the included trials are either standard care, e.g., individual care only or other types of educational programs, e.g., antenatal education programs with a smaller intervention dose than the experimental condition. In cases where two programs were compared, the most intensive was considered the experimental intervention. Co-interventions were allowed only if the intervention was delivered equally in both the experimental and control arm.

Outcome measures

We included trials reporting quantitative outcome data. Outcome data from registers, self-report, or data reported by health professionals were accepted. In trials where outcomes were measured more than one time during follow-up, we have used the measurements shortly after the intervention ends and at the longest relevant follow-up to consider the intervention effect.

In trials where an outcome was measured by the same measurement tool at the same time point and reported both as a dichotomized result (RR) and as mean of scale, we have chosen to report the mean difference as the outcome.

The primary outcomes are as follows:

- Pain relief during labor.
- Obstetric interventions.
- Psychological and social adjustment to parenthood.
- Antenatal and postnatal depression and anxiety.

The secondary outcomes are as follows:

- Knowledge acquisition.
- Maternal sense of control/active decision-making during labor and birth.
- Partner involvement at birth.
- Breast feeding success.
- Infant care abilities.
- Social support.
- Relationship satisfaction.
- Divorce/separation.

Study selection and data extraction

We conducted the selection of studies in two steps. First, two of three review authors (CSB, SFAX, and VK) independently performed the initial screening of all titles and abstracts to determine eligibility of all studies identified through the literature search. Next, two of three review authors (CSB, SGL, and VK) independently assessed the full papers identified as meeting inclusion criteria or where definite decision on exclusion could not be made from screening titles and abstracts. Any discrepancies between the assessors were resolved through discussion. A flow diagram of the selection process is shown in Additional file 3.

In some trials, the experimental and control condition received the exact same dose of antenatal education in small classes. These trials were excluded due to the difficulty of assessing the effect of antenatal education in small classes as an experimental condition as only the content varied between the experimental and the control condition.

Trials in which the experimental group received home visits, extra individual sessions, or presents for achieving the outcome in addition to the antenatal education classes were excluded as these co-interventions might have influenced the effect of the intervention beyond the effect of the classes. Extra written material to the experimental group was accepted. In trials where the intervention was 'boosted' by later individual consultations, we have used the measurement shortly before the individual consultation to consider the effect.

In cases where the content of the experimental or control condition was unclear or information incomplete, we contacted the first author by e-mail. We contacted 19 authors and received supplementary information from six of these.

Data from the included trials were extracted to summary tables containing information on the following: study design, inclusion and exclusion criteria, description of the experimental and control conditions, and outcomes of interest to the review.

Risk of bias assessment

Two review authors (CSB and VK) independently assessed the included trials according to a predefined risk of bias scoring key [8] in order to determine the

likely presence or absence of biases which might have affected the internal validity of the trials. Any discrepancies were resolved through discussion.

The scoring key includes the following characteristics:

- *Selection bias*: randomization sequence generation and allocation concealment.
- *Performance bias*: assessment of blinding of participants, educators, and outcome assessors. In trials where both subjective and objective outcomes are reported, we assessed blinding of outcome assessors separately for subjective and objective outcomes.
- *Incomplete outcome data*: assessment of systematic differences in withdrawal of study participants between the groups compared. In trials where both subjective and objective outcomes were reported, we assessed reporting bias separately for subjective and objective outcomes.
- *Selective outcome reporting bias*: assessment of systematic differences between reported and unreported findings. It was assessed whether a trial protocol exists and whether outcomes in the published trial had been reported in a pre-specified way.
- *Other sources of bias*: We assessed whether the trial was free of other sources of bias (e.g., baseline imbalance, recall bias).

First, each trial was evaluated according to each of the above-mentioned bias domains as either 'low', 'unclear', or 'high risk of bias'. Secondly, the trials were rated by an overall risk of bias. All trials rated as 'low risk of bias' in all domains were scored 'overall low risk of bias'. All other trials were scored 'overall high risk of bias'. Due to the nature of the intervention, we expected a high level of bias for the domain 'blinding of participants and educators' as it is often not possible to blind participants and educators. If all trial bias domains were rated as 'low risk of bias' with the exception of 'blinding of participants and educators', the trial was categorized as overall 'moderate risk of bias'.

'Risk of bias' tables, 'risk of bias summary', and 'risk of bias graph' for the included trials are shown in Additional file 4.

Evidence synthesis

Structured summaries of the included trials are presented in 'Characteristics of included trials' in Additional file 5. Intervention effects from the included trials are calculated and presented as risk ratios (for dichotomous outcomes) or mean differences (for continuous outcomes) with 95% confidence intervals and two-sided *P* values for each outcome and reported in effect tables [see Additional file 6] and as forest plots [see Additional file 7]. A meta-analysis was planned beforehand [9]. However, due to diverse content of experimental as well as control conditions, this was not possible to perform.

Protocol modifications

In addition to the pre-specified outcomes reported in the protocol [9], we have added relationship satisfaction and divorce/separation as secondary outcomes as these outcomes are also of great relevance as psycho-social dimensions of becoming parents.

We have reported mean differences as effect measures for continuous outcomes instead of standardized mean differences as defined by the protocol. In the process of conducting the systematic review, we found that meta-analyses could not be performed. Therefore, mean differences were preferred.

In the risk of bias assessment tool, we have included the score 'overall moderate risk of bias' for trials free of all other bias other than blinding of participants and educators and assessed the risk of bias separately for objective and self-reported outcomes.

Results

Description of included trials

We identified 5,708 records from the literature searches and an additional ten records were identified from reference lists and contact to author. A detailed flow diagram of the study selection process is shown in Additional file 3. In total, we included 17 trials in the review. We have provided full details of the included trials in the 'Characteristics of included trials' table [see Additional file 5]. A list of excluded trials with brief explanation of reasons is reported in Additional file 8.

Some trials were reported in more than one report. The 17 trials were reported in 21 papers and 1 oral presentation. Only papers fulfilling the requirements for analysis are included. The trial by Maimburg et al. was reported in two papers and one oral presentation [12-14], and the trial by Werner was reported in three papers [15-17]. For the remainder of the review, only the main report for each included trial is cited.

Results from the included trials were reported between 1988 and 2014 in 20 papers and 1 oral presentation (obtained from the author). Six trials were conducted in the United States [18-23], four trials were conducted in Australia [24-27], two were conducted in Denmark [12,16], one in the United Kingdom [28], one in Canada [29], one in Finland [30], one in Mexico [31], and one multicenter trial was conducted in Spain and France [32]. In total, the trials included 6,507 randomized women and 961 men, with a range from 74 to 1,193 participants per trial.

All trials tested the effect of antenatal education in small classes; however, the content and form of the experimental condition varied between the trials. The amount of education in the experimental condition varied from a single 1-h session [24] to 24 sessions each lasting 2.5 h [22]. Some trials focused on prevention of a specific condition among participants at specific risk, e.g., women at high risk of postnatal depression [21,28,31] or women with low socio-economic status [19,32]. Other interventions were targeted at a broader population group, e.g., all *primipara* at a specific birth site [12]. Also, control conditions differed between trials. In most of the trials, the control group was offered standard care which varied by content and amount, e.g., individual consultations with a midwife that also the experimental condition was offered [12,31]. In four trials, the control group was offered other interventions other than antenatal classes, e.g., one-to-one contact with a medical doctor [19].

Two trials were directed towards expecting fathers [23,26], and three trials specifically addressed the couple as a unit [18,22,32]. The remainders of the trials were directed towards the pregnant women, but in some of them, the expecting fathers were welcome to join one or all sessions.

For three of the pre-specified outcomes, maternal sense of control/active decision-making during labor and birth, partner involvement at birth, and infant care abilities, no data were reported. Data on pain relief during labor, obstetric interventions, knowledge acquisition, breast feeding, social support, relationship quality and divorce/separation, and psychological and social adjustment to parenthood were reported. Within the overarching category of psychological and social adjustment to parenthood, the following outcomes have been reported: antenatal and postnatal depression, anxiety, readiness for delivery and child care, self-efficacy and locus of control, co-parenting, and parent-child interaction.

Risk of bias in included trials

We assessed the risk of bias in the 17 included trials. Full details on the risk of bias scoring can be found in the 'risk of bias tables', 'risk of bias summary', and 'risk of bias graph' [see Additional file 4]. All trials except for two [27,30] reported self-reported outcomes, and two trials additionally reported objective outcomes [12,16]. Blinding of participants was only possible in one trial [32].

All trials were scored overall 'high risk of bias' for the self-reported outcomes. For the objective outcomes, two trials were scored 'overall moderate risk of bias' [12,16]. These two trials were scored 'overall high risk of bias' for the self-reported outcomes since participants were not blinded. Also, the trial by Ickovics et al. was scored 'overall high risk of bias' for the same reason although this trial had 'low risk of bias' in all other domains but reported no objective outcomes [20].

Effects of interventions

Depression prevention classes versus standard care

Three trials compared a depression-preventive program in small classes with standard care [21,28,31]. Brugha et al. examined the effect of a depression prevention antenatal program for women at risk of depression and found no significant effect on depression measured with several different measurement tools, self-efficacy, or locus of control [28]. A trial conducted by Lara et al. examined effects of a psycho-educational antenatal program among women at high risk of depression and reported no effect on depressive symptoms 6 weeks postnatally [31]. Also, Le et al. reported no effect of a psycho-educational antenatal program among women at high risk of depression - neither in pregnancy nor 6 weeks postnatally [21]. All three trials were scored 'overall high risk of bias'.

Psycho-social prevention program versus brochure on child care

One trial assessed the effect of a psycho-social prevention program for couples, compared to a brochure on child care delivered to participants in the control condition [18] on depressive symptoms, co-parenting, anxiety, and parent-child interaction for both mothers and fathers 6 months postnatally. They reported that fathers, but not mothers, in the experimental group experienced significantly higher co-parental support (MD 0.29, 0.05 to 0.53), parenting-based closeness (MD 0.35, 0.04 to 0.66), and significantly lower father-child dysfunctional interaction (MD−0.26, −0.43 to −0.09) compared to fathers in the control condition [18]. This trial was scored 'overall high risk of bias'.

Psycho-educational classes versus letter on fear of childbirth

One trial by Rouhe et al. compared the effect a group-based psycho-educational intervention directed towards women with severe fear of childbirth to written information in the form of a letter addressing fear of childbirth delivered to the participants in the control condition [30]. They found that the intervention significantly increased the likelihood of spontaneous vaginal delivery (RR 1.33, 1.11 to 1.61). They reported no effect on the use of epidural analgesia, overall caesarean section, elective and emergency caesarean section, vacuum extraction, and induction of labor [30]. This trial was scored 'overall high risk of bias'.

Program using a psycho-somatic approach versus standard antenatal education program

Ortiz Collado et al. examined the effect of an antenatal psychosomatic program designed to decrease depression

among women at high risk of postnatal depression compared to standard care [32]. They reported no significant effect on depression, social support, or relationship satisfaction among women. They also assessed relationship satisfaction among men and reported no significant effect [32]. This trial was scored 'overall high risk of bias'.

Couple-focused classes versus standard care

One trial by Schulz et al. assessed the effect of a couple-focused intervention compared to standard care on marital satisfaction among both mothers and fathers 6 months and 5.5 years postnatally as well as divorce/separation 5.5 years postnatally. They reported no significant intervention effects on any of these outcomes [22]. This trial was scored 'overall high risk of bias'.

Self-hypnosis classes versus standard care

Werner et al. compared a self-hypnosis intervention with standard care and reported no effect on the outcomes: use of epidural analgesia as pain relief during labor, spontaneous delivery, overall caesarean section, elective caesarean section, vacuum extraction, oxytocin augmentation, induction of labor, and any breast feeding 4 months postnatally [16]. However, they reported a statistically significant increased risk of emergency caesarean section (RR 1.52, 1.02 to 2.27) in the experimental group [16]. For the outcomes related to delivery, this trial was scored 'overall moderate risk of bias', while the score was 'overall high risk of bias' for breast feeding which was self-reported.

General antenatal education classes versus standard care

One trial by Maimburg et al. assessed the effect of general group-based antenatal training among primiparous compared to standard care on a range of both pharmacological and non-pharmacological pain relief outcomes, obstetric interventions, postnatal depression, breast feeding, breast feeding knowledge, and breast feeding self-efficacy [12]. They reported a protective effect on the use of epidural analgesia (RR 0.84, 0.73 to 0.98) but no significant effect on any other kind of pain relief or obstetric interventions, e.g., caesarean section and vacuum extraction. Also, no significant effects were reported on breast feeding at 5 weeks or 6 months postnatally and breast feeding self-efficacy or postnatal depression 6 weeks after birth. They reported a higher proportion with sufficient knowledge about breast feeding 6 weeks postnatally among women attending the general antenatal training program in small classes (RR 1.08, 1.01 to 1.15) [12]. For the outcomes related to delivery, this trial was scored 'overall moderate risk of bias', while the score was 'overall high risk of bias' for breast feeding, breast feeding self-efficacy, knowledge, and postnatal depression which were self-reported.

Group prenatal care (20 h) versus individual prenatal care (2 h)

A trial by Ickovics et al. examined the effect of a general antenatal education program in small classes compared to individual prenatal care (total amount of time: 2 h) [20]. They reported significantly higher scores on prenatal and infant care knowledge (MD 2.60, 1.68 to 3.52) and readiness for labor and delivery (MD 7.60, 3.34 to 11.86) at 35-weeks gestation among women in the experimental condition. They found no effect on readiness for infant care or prenatal distress at 35-weeks gestation [20]. This trial was scored 'overall high risk of bias' due to the self-report of outcomes.

Paternal education class versus standard care

Two trials examined the effect of paternal education compared to standard care [23,26]. Westney et al. conducted an intervention targeted at prospective adolescent fathers. This intervention had a significantly positive effect on paternal knowledge acquisition in relation to pregnancy, delivery, infant care, and support towards the mother (MD 9.55, 1.25 to 17.85) [23]. Maycock et al. conducted a breast feeding intervention targeted at expecting fathers. They reported a significant intervention effect on any breast feeding 6 weeks postnatally (RR 1.09, 1.00 to 1.18). There was no effect on exclusive breast feeding 6 weeks postnatally [26]. Both of these trials were scored 'overall high risk of bias'.

Extra breast feeding sessions versus standard care

In three trials, the authors examined the effect of giving extra breast feeding sessions in small classes [24,25,29]. Duffy et al. examined the effect of an antenatal group-teaching session aimed at increasing breast feeding prevalence but also reported obstetric interventions. They reported no effect on vaginal delivery, caesarean section, vacuum extraction, or forceps. They also assessed the effect on breast feeding and reported a positive effect on exclusive breast feeding 6 weeks postnatally (RR 3.20, 1.88 to 5.46) [24]. Noel-Weiss et al. examined effects of a breast feeding education workshop and reported no significant effect on breast feeding 8 weeks postnatally. However, they found a significantly higher breast feeding self-efficacy among participants in the experimental condition 4 weeks postnatally (MD 4.60, 0.72 to 8.48) but not 8 weeks postnatally [29]. Forster et al. conducted a trial comparing two breast feeding education classes with usual care. They reported no significant effect in initiation of breast feeding or breast feeding 6 months postnatally [25]. All three trials were scored 'overall high risk of bias'.

Breast feeding classes versus one-to-one contact on breast feeding

Kistin et al. assessed the effect of a breast feeding class with group discussion compared to 15- to 30-min one-to-one contact with a medical doctor on breast feeding topics and reported no effect on initiation of breast feeding or of any breast feeding 12 weeks postnatally [19]. This trial was scored 'overall high risk of bias'.

Breast feeding classes versus breast feeding and childbirth pamphlets

One trial assessed the effect of a breast feeding education program compared to breast feeding and childbirth pamphlets [27]. Rossiter reported a significantly higher rate of breast feeding initiation (RR 1.86, 1.35 to 2.55) among participant in the experimental condition but found no effect on breast feeding 6 months postnatally [27]. This trial was scored 'overall high risk of bias'.

Discussion

In this systematic review, we assessed the literature on the effect of antenatal education in small classes on obstetric and psycho-social outcomes. Across trials, the experimental and control conditions varied greatly both in their format and content, and therefore, we analyzed effect of interventions in effectively 12 different comparison groups across the 17 randomized controlled trials included. Many interventions were addressed by only one trial and the remaining in only a few trials. Due to the heterogeneity of the experimental and control conditions and outcomes, it was not appropriate to conduct meta-analysis. Most of the included trials reported on more than one outcome, and only a small number of outcomes showed statistically significant differences between the experimental and control condition. Furthermore, we found great inconsistency of results across studies, and there was no clear pattern of effect. For example, one trial assessing the effect of extra breast feeding sessions reported a positive effect on breast feeding duration [24], whereas two trials did not find an effect [25,29]. In summary, it is not possible to draw definitive conclusions on the effect of small group antenatal education on obstetric and psycho-social outcomes based on this systematic review.

Quality of the evidence

We included 17 trials. All of these were assessed as 'overall high risk of bias' for the self-reported outcomes. For the objective outcomes, two trials were scored 'overall moderate risk of bias' [12,16]. The internal validity of the results of this review is therefore limited. Also, generally sample sizes were small - 12 of the 17 trials were conducted with fewer than 400 individuals randomized. There was a tendency that the larger and more recent trials had fewer methodological limitations and were reported in more detail than the earlier trials with smaller sample sizes. There is a need for trial authors to report trials according to the CONSORT principles [33].

Strengths and limitations

We used the Cochrane Handbook for Systematic Reviews of Interventions [8] as a guide for conducting this systematic review. We registered the review within the International Prospective Register of Systematic Reviews (PROSPERO) and published our methods as a protocol before conducting the review [9]. We conducted a thorough literature search performed by an information specialist and had no restrictions regarding language and publication date. Two review authors independently extracted data and scored risk of bias according to a detailed bias assessment tool.

The trials included in this review are very diverse regarding experimental conditions, control conditions, and populations studied and are therefore difficult to compare. The strength of the conclusions is limited by sparse and lower quality of evidence.

In 2007, a systematic review by Gagnon and Sandall was conducted [3] evaluating the effect of both individual and group antenatal education for childbirth or parenthood. They concluded that high-quality evidence was lacking and that the effects of antenatal education are largely unknown. In this review, we specifically focused on antenatal education in small classes conducted in a Western setting and assessed the literature up to 2014. Also, in the present review, we found limited evidence from which to draw conclusions regarding the effect of antenatal education in small classes. We chose to focus primarily on evaluating evidence about the form of antenatal education, i.e., education in small classes and not the content as such. We excluded trials evaluating two programs with the same dose of antenatal education in small classes. To look into the effect of content, it would be relevant to conduct a systematic review evaluating this aspect.

Implications for research

There is a need to conduct high-quality, randomized trials with adequate sample sizes and transparent reporting of relevant outcome measures to evaluate the effect of antenatal education in small classes. Results from a large ongoing randomized trial will soon be available [34]. Given the uncertainty in effects and costs of small group antenatal education, we would recommend that future trials should first focus on a comparison to standard care rather than comparing the relative effects of different educational programs. Future trials should also initially assess the feasibility of interventions in order that they develop and evaluate educational programs that are

likely to be implementable in an everyday clinical practice setting, if proven effective. Finally, there is the issue of the trial population and whether to apply the educational intervention to the broad population or to limit it to high-risk populations, such as those with depression.

Implications for practice

No clear recommendations for practice can be made based on the results of this review. The trials included all varied greatly in extent, method, and content, and a meta-analysis was not possible to perform. This makes it difficult to compare results across trials.

Conclusions

Insufficient evidence exists as to whether antenatal education in small classes has any effect on obstetric or psycho-social outcomes. Given that the evidence base is inconclusive, emerging evidence from future well-conducted and well-reported trials may help to make conclusions about the effectiveness of antenatal education in small classes. We recommend updating this review regularly with emerging evidence.

Additional files

Additional file 1: PRISMA 2009 checklist. The file contains a filled in PRISMA checklist for the systematic review.

Additional file 2: Search strategy. The file contains the search strategy used in the databases Medline, EMBASE, CENTRAL, CINAHL, Web of Science, and PsycINFO.

Additional file 3: PRISMA 2009 Flow Diagram. The file contains a flow diagram of the trial selection process.

Additional file 4: Risk of bias tables. The file contains the assessment of risk of bias for each included trial, a risk of bias summary, a risk of bias graph.

Additional file 5: Characteristics of included trials. The file contains characteristics of design, participants, content of experimental and control conditions, and outcomes for each included trial.

Additional file 6: Effect tables. The file contains tables of measures of intervention effects (RR and MD) with 95% confidence intervals and two-sided *P* values for each outcome in the included trials.

Additional file 7: Forest plots. The file contains forest plots of intervention effects (RR and MD) for each outcome in the included trials.

Additional file 8: Characteristics of excluded trials. The file contains a list of excluded trials with brief explanations of reasons for exclusion.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

CSB, SFA, SGL, PD, and VK developed the design of the systematic review and drafted the manuscript. VK, CSB, and SGL extracted data. VK and CSB scored the risk of bias. SKA developed the search strategy and performed the literature search. All authors have read and approved the manuscript.

Authors' information

CSB: Master of Science in Public Health, PhD student on a large randomized trial (the NEWBORN trial) evaluating the effect of a structured antenatal education program.
SFA: Pharmaconomist, Exam.pharm.cons, MPH; research assistant on the NEWBORN trial.
SGL: Bachelor student in Public Health, student assistant on the NEWBORN trial.
SKA: Information specialist. Has expertise in literature searching.
PD: Professor, Dr.Med.Sci.; workpackage-chair of Child Intervention Research as part of The Centre of Intervention Research and research director of the research program for Child and Adolescent Health at The National Institute of Public Health, University of Southern Denmark.
VK: Midwife, MPH, PhD; principal investigator of the NEWBORN trial.

Acknowledgements

We would like to thank Jane Lindschou, Clinical Research Consultant, Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Denmark, for her valuable input in designing this systematic review.
Also, we thank Rod Taylor, Visiting Professor, National Institute of Public Health, University of Southern Denmark for his valuable input during the revision of this systematic review.
The NEWBORN trial is funded by the Danish Cancer Society, and this systematic review is related to the NEWBORN trial. The Danish Cancer Society does not have any influence on the decision to submit this manuscript.

Received: 15 July 2014 Accepted: 3 February 2015
Published online: 28 February 2015

References

1. Ahlden I, Ahlehagen S, Dahlgren LO, Josefsson A. Parents' expectations about participating in antenatal parenthood education classes. J Perinat Educ. 2012;21:11–7.
2. Fabian HM, Radestad IJ, Waldenstrom U. Childbirth and parenthood education classes in Sweden. Women's opinion and possible outcomes. Acta Obstet Gynecol Scand. 2005;84:436–43.
3. Gagnon AJ, Sandall J. Individual or group antenatal education for childbirth or parenthood, or both. Cochrane Database Syst Rev. 2007;3, CD002869.
4. Illeris KJP, Jarvis P, Kegan R, Engeström Y, Elkjaer B, Mezirow J, et al. Contemporary theories of learning: learning theorists…in their own words. New York: Routledge; 2009.
5. Ip WY, Tang CS, Goggins WB. An educational intervention to improve women's ability to cope with childbirth. J Clin Nurs. 2009;18:2125–35.
6. Moat KA, Lavis JN, Wilson MG, Rottingen JA, Barnighausen T. Twelve myths about systematic reviews for health system policymaking rebutted. J Health Serv Res Policy. 2013;18:44–50.
7. Janicke DM, Finney JW. Children's primary health care services: social-cognitive factors related to utilization. J Pediatr Psychol. 2003;28:547–57.
8. Higgins JPT, Green S. Cochrane Handbook for Systematic Reviews of Interventions. West Sussex, England: The Cochrane Collaboration and John Wiley & Sons Ltd.; 2011.
9. Brixval CS, Axelsen SF, Andersen SK, Due P, Koushede V. The effect of antenatal education in small classes on obstetric and psycho-social outcomes: a systematic review and meta-analysis protocol. Syst Rev. 2014;3:12.
10. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6:e1000097.
11. Members and partners of OECD. [http://www.oecd.org/about/membersandpartners/]
12. Maimburg RD, Vaeth M, Durr J, Hvidman L, Olsen J. Randomised trial of structured antenatal training sessions to improve the birth process. BJOG. 2010;117:921–8.
13. Kronborg H, Maimburg RD, Vaeth M. Antenatal training to improve breast feeding: a randomised trial. Midwifery. 2012;28:784–90.
14. Maimburg RD. Postpartum Depression - Results and Characteristics from a randomised trial. Oral presentation at the Nordic Midwifery Congress, Olso, Norway, 14 June, 2013.
15. Werner A, Uldbjerg N, Zachariae R, Nohr EA. Effect of self-hypnosis on duration of labor and maternal and neonatal outcomes: a randomized controlled trial. Acta Obstet Gynecol Scand. 2013;92:816–23.

16. Werner A, Uldbjerg N, Zachariae R, Rosen G, Nohr EA. Self-hypnosis for coping with labour pain: a randomised controlled trial. BJOG. 2013;120:346–53.
17. Werner A, Uldbjerg N, Zachariae R, Wu CS, Nohr EA. Antenatal hypnosis training and childbirth experience: a randomized controlled trial. Birth (Berkeley, Calif). 2013;40:272–80.
18. Feinberg ME, Kan ML. Establishing family foundations: intervention effects on coparenting, parent/infant well-being, and parent-child relations. J Fam Psychol. 2008;22:253–63.
19. Kistin N, Benton D, Rao S, Sullivan M. Breast-feeding rates among black urban low-income women: effect of prenatal education. Pediatrics. 1990;86:741–6.
20. Ickovics JR, Kershaw TS, Westdahl C, Magriples U, Massey Z, Reynolds H, et al. Group prenatal care and perinatal outcomes: a randomized controlled trial. Obstet Gynecol. 2007;110:330–9.
21. Le HN, Perry DF, Stuart EA. Randomized controlled trial of a preventive intervention for perinatal depression in high-risk Latinas. J Consult Clin Psychol. 2011;79:135–41.
22. Schulz MS, Cowan CP, Cowan PA. Promoting healthy beginnings: a randomized controlled trial of a preventive intervention to preserve marital quality during the transition to parenthood. J Consult Clin Psychol. 2006;74:20–31.
23. Westney OE, Cole OJ, Munford TL. The effects of prenatal education intervention on unwed prospective adolescent fathers. J Adolesc Health Care. 1988;9:214–8.
24. Duffy EP, Percival P, Kershaw E. Positive effects of an antenatal group teaching session on postnatal nipple pain, nipple trauma and breast feeding rates. Midwifery. 1997;13:189–96.
25. Forster D, McLachlan H, Lumley J, Beanland C, Waldenstrom U, Amir L. Two mid-pregnancy interventions to increase the initiation and duration of breastfeeding: a randomized controlled trial. Birth (Berkeley, Calif). 2004;31:176–82.
26. Maycock B, Binns CW, Dhaliwal S, Tohotoa J, Hauck Y, Burns S, et al. Education and support for fathers improves breastfeeding rates: a randomized controlled trial. J Hum Lact. 2013;29:484–90.
27. Rossiter JC. The effect of a culture-specific education program to promote breastfeeding among Vietnamese women in Sydney. Int J Nurs Stud. 1994;31:369–79.
28. Brugha TS, Wheatley S, Taub NA, Culverwell A, Friedman T, Kirwan P, et al. Pragmatic randomized trial of antenatal intervention to prevent post-natal depression by reducing psychosocial risk factors. Psychol Med. 2000;30:1273–81.
29. Noel-Weiss J, Rupp A, Cragg B, Bassett V, Woodend AK. Randomized controlled trial to determine effects of prenatal breastfeeding workshop on maternal breastfeeding self-efficacy and breastfeeding duration. J Obstet Gynecol Neonatal Nurs. 2006;35:616–24.
30. Rouhe H, Salmela-Aro K, Toivanen R, Tokola M, Halmesmaki E, Saisto T. Obstetric outcome after intervention for severe fear of childbirth in nulliparous women - randomised trial. BJOG. 2013;120:75–84.
31. Lara MA, Navarro C, Navarrete L. Outcome results of a psycho-educational intervention in pregnancy to prevent PPD: a randomized control trial. J Affect Disord. 2010;122:109–17.
32. Ortiz Collado MA, Saez M, Favrod J, Hatem M. Antenatal psychosomatic programming to reduce postpartum depression risk and improve childbirth outcomes: a randomized controlled trial in Spain and France. BMC Pregnancy Childbirth. 2014;14:22.
33. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ (Clin Res Ed). 2010;340:c332.
34. Koushede V, Brixval CS, Axelsen SF, Lindschou J, Winkel P, Maimburg RD, et al. Group-based antenatal birth and parent preparation for improving birth outcomes and parenting resources: study protocol for a randomised trial. Sex Reprod Healthc. 2013;4:121–6.

Additional files

PRISMA checklist

Section/topic	#	Checklist item	Reported on page #*
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	Title page 1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	3-4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	4
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	4
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	4
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	Additional file 2
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6 + flow diagram In additional file 3
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	6
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	7
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	7 + risk of bias tables in additional file 4
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	7
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	Not relevant

Search strategy

Web of Science

All indexes, all years. Filters: none.

1: TS=(antenatal OR prenatal OR pregnancy OR birth OR childbirth OR (labor OR labour) OR obstetric OR (delivery OR deliveries))
2: TS=(education OR "parent education*" OR preparation OR "parent preparation" OR "early intervention")
3: TS=("randomi* control* trial*" OR "randomi* trial*" OR "randomi* clinical trial*")
4: 1 AND 2 AND 3.

Medline

Filters: Refined by *randomized controlled trial*, *humans*.

1: TS=(antenatal OR prenatal OR pregnancy OR birth OR childbirth OR (labor OR labour) OR obstetric OR (delivery OR deliveries)) 2: TS=(education OR "parent education*" OR preparation OR "parent preparation" OR "early intervention") 3: 1 AND 2

Cinahl

Filters: Refined by *randomized controlled trial*.

1: SU= antenatal OR prenatal OR pregnancy OR birth OR childbirth OR (labor OR labour) OR obstetric OR (delivery OR deliveries))
2: SU=education OR "parent education*" OR preparation OR "parent preparation" OR "early intervention"
3: 1 AND 2

Additional Cinahl search with no filters:

4: SU=antenatal OR prenatal OR pregnancy OR birth OR childbirth OR (labor OR labour) OR obstetric OR (delivery OR deliveries)
5: SU=education OR "parent education*" OR preparation OR "parent preparation" OR "early intervention")
6: SU="randomi* control* trial*" OR "randomi* trial*" OR "randomi* clinical trial*"
7: 4 AND 5 AND 6
8: 3 AND 7

Cochrane

Filter: Title, abstract, keyword, refined by *trials*.

1: antenatal OR prenatal OR pregnancy OR birth OR childbirth OR (labor OR labour) OR obstetric OR (delivery OR deliveries))
2: (education OR "parent education*" OR preparation OR "parent preparation" OR "early intervention")
3: ("randomi* control* trial*" OR "randomi* trial*" OR "randomi* clinical trial*")
4: 1 AND 2 AND 3

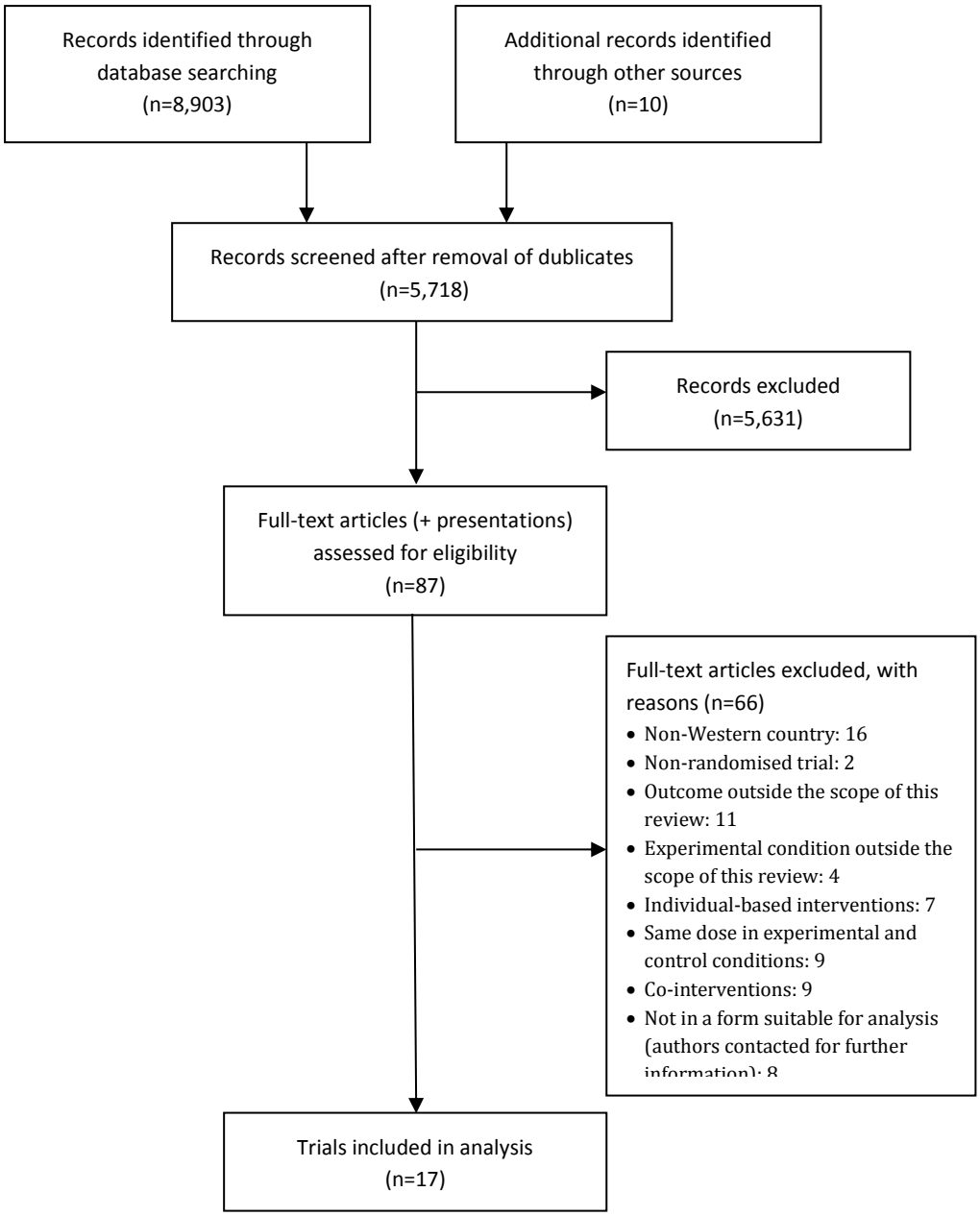
Embase / Psycinfo

Embase and Psycinfo were searched together in the same database. No filters.

In total, four combined searches were made; in abstracts (AB), in keywords (key), in subject headings (SH), and in titles (TI).

1: AB: antenatal OR prenatal OR pregnancy OR birth OR childbirth OR (labor OR labour) OR obstetric OR (delivery OR deliveries)
2: AB: education OR "parent education*" OR preparation OR "parent preparation" OR "early intervention"
3 ALL FIELDS: "randomi* control* trial*" OR "randomi* trial*" OR "randomi* clinical trial*"
4: 1 AND 2 AND 3
5: KEY: antenatal OR prenatal OR pregnancy OR birth OR childbirth OR (labor OR labour) OR obstetric OR (delivery OR deliveries)
6: KEY: education OR "parent education*" OR preparation OR "parent preparation" OR "early intervention"
7: ALL FIELDS: "randomi* control* trial*" OR "randomi* trial*" OR "randomi* clinical trial*"
8: 5 AND 6 AND 7
9: SH: antenatal OR prenatal OR pregnancy OR birth OR childbirth OR (labor OR labour) OR obstetric OR (delivery OR deliveries)
10: SH: education OR "parent education*" OR preparation OR "parent preparation" OR "early intervention"
11: ALL FIELDS: "randomi* control* trial*" OR "randomi* trial*" OR "randomi* clinical trial*"
12: 9 AND 10 AND 11
13: TI: antenatal OR prenatal OR pregnancy OR birth OR childbirth OR (labor OR labour) OR obstetric OR (delivery OR deliveries)
14: TI: education OR "parent education*" OR preparation OR "parent preparation" OR "early intervention"
15: ALL FIELDS: "randomi* control* trial*" OR "randomi* trial*" OR "randomi* clinical trial*"
16: 13 AND 14 AND 15
17: 4 AND 8 AND 12 AND 16

PRISMA Flow diagram



Risk of bias

Brugha 2000, Overall risk of bias: high

Risk of bias	Description	Judgement
Sequence generation	Computerized stratified randomization, using minimization on three prognostic factors: level of social support, screen GHQ-D and ethnic group, was used by the research interviewer to allocate half the consenting women to intervention and half to control	Low
Allocation concealment	No description.	Unclear
Blinding of participants and educators	Not possible to blind participants. The educators were not blinded but "…they were not involved in any way in intervention allocation or in the research assessments".	High
Blinding of outcome assessor (self-reported outcomes)	Not possible to blind outcome assessors (self-reported outcomes). Outcome assessors (interviewers) were blinded towards intervention group, and the women were asked not to reveal their status. "The allocation code was not broken until completion of the fieldwork and primary analyses." "Analysis of the interviewer's records of which group she thought each respondent had been allocated to showed no difference from chance."	High
Incomplete outcome data (self-reported outcomes)	Nonresponse rate was 9 % in both the experimental and control condition. Sensitivity analyses testing the influence of these missing outcome data did not alter the results.	Low
Selective reporting bias	No study protocol found. All stated primary and secondary outcome measures stated in the paper are reported.	Low
Other sources of bias		Low

Duffy 1997, Overall risk of bias: high

Risk of bias	Description	Judgement
Sequence generation	Randomized in blocks of 12 – no further description of sequence generation.	Unclear
Allocation concealment	Group allocation was blinded to the researcher. "Randomization was achieved using a sealed envelope technique….".	Low
Blinding of participants and educators	Not possible to blind participants or educators.	High
Blinding of outcome assessor (self-reported outcomes)	Not possible to blind outcome assessors (self-reported outcomes).	High
Incomplete outcome data (self-reported outcomes)	Two women in the experimental group were excluded from data analysis due to revelation of treatment condition. 3 women in the control condition excluded. No information on differences in characteristics. Full response rate on the remaining – 35 participants in each condition.	Low
Selective reporting bias	No study protocol found, but all listed primary outcomes in the paper are reported.	Unclear
Other sources of bias		Unclear

Feinberg 2008, Overall risk of bias: high

Risk of bias	Description	Judgement
Sequence generation	No description of procedure	Unclear
Allocation concealment	No description	Unclear
Blinding of participants and educators	Not possible to blind participants or educators.	High
Blinding of outcome assessor (self-reported outcomes)	Not possible to blind outcome assessors (self-reported outcomes).	High
Incomplete outcome data (self-reported outcomes)	No differential drop-out rates at 6 months follow-up; 11 % in the experimental and 9 % in the control condition. There was no evidence of differential attrition by condition.	Low
Selective reporting bias	Reporting of results in 3 papers. No indication of relevant outcomes not reported.	Low
Other sources of bias		Unclear

Forster 2004, Overall risk of bias: high

Risk of bias	Description	Judgement
Sequence generation	"A computerized system of biased urn randomization"	Low
llocation concealment	Randomization "was accessed by telephone by the research midwife to ascertain women's group allocation".	Low
Blinding of participants and educators	Not possible to blind participants or educators.	High
Blinding of outcome assessor (self-reported outcomes)	Not possible to blind outcome assessors (self-reported outcomes).	High
Incomplete outcome data (self-reported outcomes)	Non-response rate for experimental group: 5 % at 2-4 days after birth and 11 % at 6 months. Control group: 5 % at 2-4 days after birth and 9 % at 6 months. No reporting on differences in characteristics of non-responders between groups.	Unclear
Selective reporting bias	According to study protocol all listed primary outcomes are reported.	Low
Other sources of bias	35 women were unable to be interviewed at first follow-up. They answered questions later – may cause recall bias. No information on differences in traceability of respondents between conditions given.	Unclear

Ickovicks 2007, Overall risk of bias: high

Risk of bias	Description	Judgement
Sequence generation	"A computer-generated randomization sequence, password protected to recruitment staff and participants, was used to assign participants".	Low
Allocation concealment	"Allocation was concealed from participant and research staff until eligibility screening was completed and study condition was assigned. These tasks were completed by trained research team members who were independent of prenatal care".	Low
Blinding of participants and educators	Not possible to blind patients and educators.	High
Blinding of outcome assessor (self-reported outcomes)	Not possible to blind outcome assessors (self-reported outcomes).	High
Incomplete outcome data (self-reported outcomes)	No differential drop-out between experimental (10 %) and control group (11.5 %) in the measurements in week 35. No reported drop-out number for breast-feeding initiation.	Low
Selective reporting bias	No study protocol found, but all listed primary outcomes in the paper are reported.	Low
Other sources of bias	There were differences in some of the baseline characteristics. Authors made analyses adjusted for these variables. This did not change the overall significance level.	Low

Kistin 1990, Overall risk of bias: high

Risk of bias	Description	Judgement
Sequence generation	"Women… were randomly, using a random numbers table,…"	Low
Allocation concealment	No description.	Unclear
Blinding of participants and educators	Not possible to blind patients and educators.	High
Blinding of outcome assessor (self-reported outcomes)	Not possible to blind outcome assessors (self-reported outcomes).	High
Incomplete outcome data (self-reported outcomes)	From the information given it is not possible to calculate drop-out rates for each of the two groups included in the review. Overall, of the 159 women who agreed to participate, 29 dropped out (18 %). Drop-outs differed in age and breast feeding plans (not significant).	High
Selective reporting bias	No study protocol found. No outcomes other than the reported are listed.	Unclear
Other sources of bias	Large differences in some of the baseline characteristics related to the outcome.	High

Lara 2010, Overall risk of bias: high

Risk of bias	Description	Judgement
Sequence generation	Randomization was performed using a blocked randomization procedure. Blocks were sequentially opened every time a group started. To ensure conditions were balanced, an envelope contained equal number of folded papers for each one for the first two groups.	Low
Allocation concealment	No description	High
Blinding of participants and educators	Not possible to blind participants or educators.	High
Blinding of outcome assessor (self-reported outcomes)	Not possible to blind outcome assessors (self-reported outcomes).	High
Incomplete outcome data (self-reported outcomes)	Non-response rate were much higher in the experimental group (72 %) 6 weeks after birth than in the control group (39 %). No reporting on differences in characteristics of non-responders between groups.	High
Selective reporting bias	No study protocol found. No other outcomes are listed in the paper.	Unclear
Other sources of bias	Women in the intervention group had higher rates of depressive symptoms and anxiety at baseline.	High

Le 2011, Overall risk of bias: high

Risk of bias	Description	Judgement
Sequence generation	No description	Unclear
Allocation concealment	"… a sealed envelope with her group membership was assigned by the first author; neither participant nor interviewer knew the result of the random assignment until this envelope was opened."	Low
Blinding of participants and educators	Not possible to blind participants or educators.	High
Blinding of outcome assessor (self-reported outcomes)	Not possible to blind outcome assessors (self-reported outcomes). Interviewers at each outcome assessment time point were not blind to group assignment.	High
Incomplete outcome data (self-reported outcomes)	Nonresponse rate at the post intervention assessment was 16 % in the experimental group and 12 % in the control condition. At 6 weeks postpartum, the nonresponse rate was higher in the experimental group (21 %) than in the control group (13 %). No data on differences in characteristics are presented.	High
Selective reporting bias	No study protocol found. No other outcomes are listed in the paper.	Unclear
Other sources of bias		Unclear

Maimburg 2010, Overall risk of bias: *Objective outcomes: moderate; Self-reported outcomes: high*

Risk of bias	Description	Judgement
Sequence generation	"The randomization program used an algorithm generated by a data manager".	Low
Allocation concealment	"Randomization was assigned by a staff midwife using a computer-assisted voice response system"	Low
Blinding of participants and educators	Not possible to blind participants and educators.	High
Blinding of outcome assessor		
Objective outcomes	Outcome assessors (midwives in the maternity ward) were blinded towards intervention group.	Low
Self-reported outcomes	Not possible to blind outcome assessors (self-reported outcomes).	High
Incomplete outcome data		
Objective outcomes	In both the experimental and control group non-report on obstetric outcomes were 3 %.	Low
Self-reported outcomes	In the measure of breast-feeding at 6 weeks the non-response rate for both groups were around 10 %. No data on differences in characteristics are presented.	Unclear
Selective reporting bias	All main outcomes reported in study protocol are reported in the main paper of the study. The secondary outcomes are reported in other papers or oral presentations.	Low
Other sources of bias	Sample size calculations based on the primary outcome.	Low

Maycock 2013, Overall risk of bias: high

Risk of bias	Description	Judgement
Sequence generation	Randomization with a random number generator.	Low
Allocation concealment	Participants were randomized, with no blinding	High
Blinding of participants and educators	Not possible to blind participants or educators.	High
Blinding of outcome assessor (self-reported outcomes)	Not possible to blind outcome assessors (self-reported outcomes).	High
Incomplete outcome data (self-reported outcomes)	Not possible to calculate non-response rates for each group separately.	Unclear
Selective reporting bias	No study protocol found. The primary outcome is reported.	Unclear
Other sources of bias	Sample size calculations based on the primary outcome.	Low

Noel-Weiss 2006, Overall risk of bias: high

Risk of bias	Description	Judgement
Sequence generation	Sequentially numbered envelopes	Unclear
Allocation concealment	"Participants returned the registration package in a sealed manila envelope, and randomization was completed by matching the manila envelope with a sealed, sequentially numbered, opaque envelope containing a slip of paper stating either Control or Workshop"	Low
Blinding of participants and educators	Not possible to blind participants or educators.	High
Blinding of outcome assessor (self-reported outcomes)	Not possible to blind outcome assessors (self-reported outcomes).	High
Incomplete outcome data (self-reported outcomes)	Not possible to calculate the non-response rate for each group separately.	Unclear
Selective reporting bias	No study protocol found, but all listed primary outcomes in the paper are reported.	Unclear
Other sources of bias		Low

Ortiz Collado 2014, Overall risk of bias: high

Risk of bias	Description	Judgement
Sequence generation	”A statistician produced a computer generated random distribution of women with antenatal risk of PPD in both groups, EG and CG”.	Low
Allocation concealment	“The allocation to the study groups was blinded; all interviews were sent to an outside statistician…. The statistician telephoned the researcher to notify the assignment of eligible women to control groups or experimental groups”.	Low
Blinding of participants and educators	“Participants knew they were in a study group but did not know the distinction between control and experimental intervention. The nurse midwives who ran the control group also had no prior knowledge. Only nurse midwives who animated the experimental group knew about the distinction but never had access to the questionnaires and never knew the evaluated variables.”	Low
Blinding of outcome assessor (self-reported outcomes)	Outcome assessors (self-reported outcomes) were blinded.	Low
Incomplete outcome data (self-reported outcomes)	Non-response rate for the questionnaire were higher in the control group (36 %) than in the experimental group (25 %). No data on differences in characteristics are presented.	High
Selective reporting bias	No study protocol found. Both the stated primary and secondary outcome stated in the paper are reported.	Unclear
Other sources of bias		Low

Rossiter 1994, Overall risk of bias: high

Risk of bias	Description	Judgement
Sequence generation	No description of procedure. Large differences in number of participants assigned to the two conditions.	High
Allocation concealment	No description.	Unclear
Blinding of participants and educators	Not possible to blind participants or educators.	High
Blinding of outcome assessor (objective outcomes)	No description	Unclear
Incomplete outcome data (objective outcomes)	Non-response rates: post-test: 6 %, 6 months following birth: 10 %. No reporting on non-response rates for experimental and control groups separately.	Unclear
Selective reporting bias	No study protocol found, but all listed primary outcomes in the paper are reported.	Unclear
Other sources of bias	Large differences in baseline characteristics related to the outcome which may have biased the results.	High

Rouhe 2012, Overall risk of bias: high

Risk of bias	Description	Judgement
Sequence generation	No description.	Unclear
Allocation concealment	".. were randomised … to the intervention or control group in the proportion of 1:2 in balanced blocks of 18 by sealed opaque envelopes."	Low
Blinding of participants and educators	Not possible to blind participants or educators.	High
Blinding of outcome assessor (objective outcomes)	The outcomes were assessed from medical records.	Low
Incomplete outcome data (objective outcomes)	Full response rate.	Low
Selective reporting bias	No study protocol found. The listed primary outcome is reported.	Unclear
Other sources of bias	Discrepancies between numbers of randomized women in the two papers reporting from the study.	High

Schulz 2006, Overall risk of bias: high

Risk of bias	Description	Judgement
Sequence generation	"Expectant couples were randomized to condition using a random number table.."	Low
Allocation concealment	No description.	Unclear
Blinding of participants and educators	Not possible to blind participants or educators.	High
Blinding of outcome assessor (self-reported outcomes)	Not possible to blind outcome assessors (self-reported outcomes).	High
Incomplete outcome data (self-reported outcomes)	94 % returned the 6 month follow-up questionnaire – 96 % in the experimental group and 92 % in the control group. At the 66 months follow-up 46 % in the experimental and 66 % in the control group returned questionnaire. No information on differences in characteristics among responders and non-responders are given. The drop-out rates at 66 months are substantial and differ between groups.	High
Selective reporting bias	No specification of other collected outcomes in the trial.	Unclear
Other sources of bias		Unclear

Werner 2013, Overall risk of bias: *Objective outcomes: moderate, Self-reported outcomes: high*

Risk of bias	Description	Judgement
Sequence generation	Computer-generated system.	Low
Allocation concealment	"The participants were randomly allocated… using a computer-generated interactive voice response telephone randomization system".	Low
Blinding of participants and educators	Not possible to blind participants or educators.	High
Blinding of outcome assessor		
Objective outcomes	Outcome assessor for the birth related outcomes (midwives assisting the birth) were blinded to the participant's allocated treatment. Outcomes were extracted from an ongoing data collection from all births at the hospital or from medical records.	Low
Self-reported outcomes	Not possible to blind outcome assessors (self-reported outcomes).	High
Incomplete outcome data		
Objective outcomes	Full response rate	Low
Self-reported outcomes	Response rate for the 6 week questionnaire were high (97 % and 98.4 % in control group and experimental group respectively). At 6 months after birth the corresponding rates were 96.1 % and 96.8 %. No data on differences in characteristics are presented, but very low drop-out rate in both conditions.	Low
Selective reporting bias	Study protocol available. The stated primary outcome is reported, and only a few secondary outcomes are not yet reported.	Low
Other sources of bias		Low

Westney 1988, Overall risk of bias: high

Risk of bias	Description	Judgement
Sequence generation	No description of randomization procedure.	Unclear
Allocation concealment	No description.	Unclear
Blinding of participants and personnel	Not possible to blind participants and educators	High
Blinding of outcome assessor (self-reported outcomes)	Not possible to blind outcome assessors (self-reported outcomes).	High
Incomplete outcome data (self-reported outcomes)	No drop-out from baseline to follow-up.	Low
Selective reporting bias	No study protocol found. No outcomes other than the reported are listed.	Unclear
Other sources of bias		Unclear

Figure 4.a. Risk of bias summary: review authors' judgments about each risk of bias item for each included study

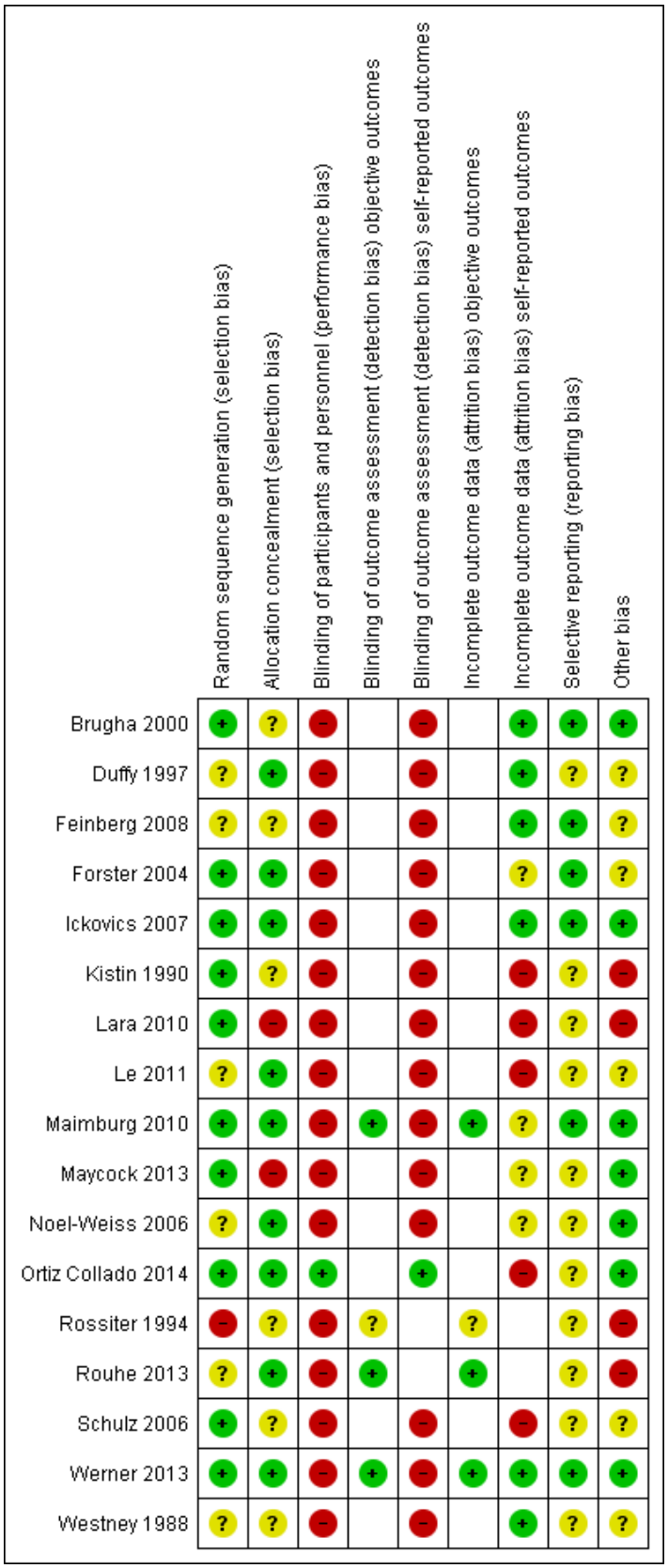
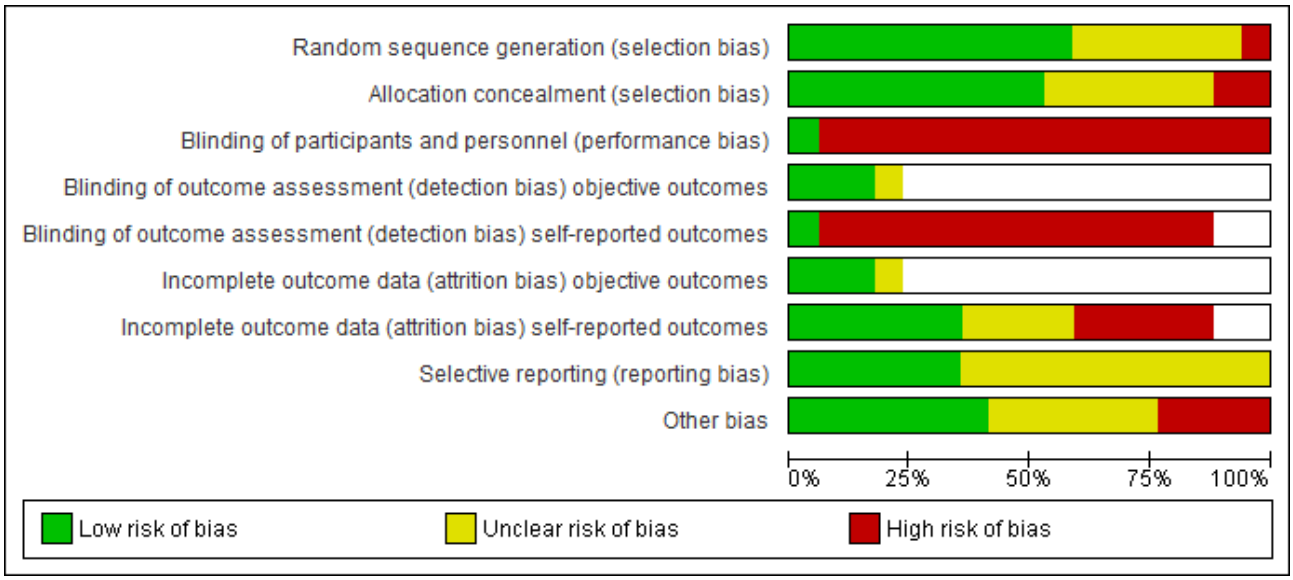


Figure 4.b. Risk of bias graph: review authors' judgments about each risk of bias item presented as percentages across all included studies



Characteristics of included trials

	Design	Participants	Experimental and control conditions	Outcomes of interest to the review
Brugha 2000	Individually randomized trial conducted in Leicester, UK.	209 women in risk of postnatal depression. Experimental (n=103), control (n=106). **Inclusion criteria:** pregnant woman screened at risk of postnatal depression, at least 16 years of age, primipara, planning to continue pregnancy to full-term, residing within reasonable travelling distance of the hospital, capable of understanding and completing screening questionnaires in English and of giving written, informed consent.	**Experimental condition**: • A depression-preventive program consisting of 6 structured 2-hour weekly antenatal classes and a postnatal reunion class. • Starting in week 28 of gestation. • Group size was 8-16 women. • General education was given rather than formal lectures – both discussion exercises and role plays were used. The group was expected to provide emotional support. The woman's partner or significant other was encouraged to attend one session. • Instructors were nurses and occupational therapists, with extensive experience in hospital and community general psychiatry. **Control condition:** Standard care consisting of 10 individual consultations with a community midwife. No specific focus on depression prevention.	All outcomes were assessed three months postnatal with self-completion questionnaires: **The modified GHQ-D.** Depression defined as two or more symptoms of depression. **The Edinburgh Postnatal Depression Scale (EPDS):** A self-administered 10-item questionnaire. A cut-off point of ≥11 was used as an indicator of risk of postnatal depression. **The Schedules for Clinical Assessment in Neuropsychiatry (SCAN):** A semi-structured clinical interview. Depression was defined according to ICD-10 criteria for depressive disorder. **Self-efficacy and locus of control:** Measured by three single items.
Duffy 1997	Individually randomized trial conducted in Western Australia. Planned as a pilot study.	75 primarily low-income women. Experimental (n=37), control (n=38). **Inclusion criteria:** primipara women who attended the antenatal clinic of the study hospital; intention to breast feed. **Exclusion criteria:** delivery before 37 weeks gestation; medical complications.	**Experimental condition:** • One additional one-hour breast feeding session. • Group size was 6 women. • Session was provided after gestation week 36. • The content of the teaching session was the correct position and attachment of the baby on the breast for feeding. • The instructor was a midwife who was also a senior lactation consultant. **Control condition:** Standard antenatal education (content not described).The experimental condition received this offer too,	**Obstetric interventions:** Spontaneous vaginal delivery, caesarean section, vacuum extraction, forceps. Measured by questionnaire 24 hours following birth. **Breast feeding duration:** measured by questionnaire six weeks postnatally. Breast feeding defined as exclusive breast feeding.
Feinberg 2008	Individually randomized trial conducted I USA.	169 primarily non-Hispanic white couples. Experimental (n=89), control (n=80). **Inclusion criteria:** primipara; living together (regardless of marital status). All participants were at least 18 years of age.	**Experimental condition:** • Psychosocial prevention program for couples with 4 prenatal and 4 postnatal sessions. • Each group consisted of 6–10 couples. • Focus was on emotional self-management, conflict management, problem solving, communication, and mutual support strategies that foster positive joint parenting of an infant. The program was manualized, with didactic material, exercises, and behavioral	All outcomes were measured by self-reported questionnaires from both mother and father six months postnatally. **Coparenting:** both parents reported on multiple dimensions (three scales) of the coparenting relationship with a measure developed for the study. All 15 items utilized 7-point Likert response scales.

			rehearsal included in the curriculum for each session. • Sessions were led by a male–female team. • No information on how long each session lasted. **Control condition:** Couples were mailed a brochure about selecting quality child care.	**Depressive symptoms:** measured with a subset of seven items from the Center for Epidemiological Studies Depression Scale. Items were answered on a 4-point frequency scale. **Anxiety:** measured with the 20-item short form of the Taylor Manifest Anxiety Scale, which measures chronic anxiety. Items were answered yes/no. **Parent–Child dysfunctional interaction:** assessed by the 6-item Dysfunctional Interaction Scale from the Parental Stress Index.
Forster 2004	Individually randomized controlled trial with 3 arms: 2 different experimental and 1 control condition. In this review effects of the most intensive intervention group (attitudes) against the control group is tested. Conducted in Melbourne, Australia.	654 relatively disadvantaged, low-income women. 92.5 % planned to breast feed. Experimental (n=327), control (n=327). **Inclusion criteria:** booking as public patients; primipara; between 16 and 24 weeks pregnant at time of recruitment; and able to speak, read, and write in English. **Exclusion criteria:** physical problems that prevented breast feeding; and choosing birth center or private obstetric care.	**Experimental condition:** • Two 1-hour breast feeding education sessions. • Participants were approximately 20 to 25 weeks' gestation. • Class size of approximately 8 women. Women were encouraged to bring their partners or a significant other. • Sessions focused on changing attitudes to breast feeding, and included information about the advantages of breast feeding, an exploration of the expectant parents' views and attitudes on breast feeding, and their perceptions of the views of their family and friends, as well as community attitudes, and group discussion. • Sessions were led by midwives and a community educator. **Control condition:** Standard care including: formal breastfeeding education sessions; lactation consultant support; community breastfeeding groups; attendance at a breastfeeding information evening; 24-hour telephone counseling support; and a postnatal home visit by a domiciliary midwife. The experimental condition received this offer too,	**Breast feeding initiation:** measured by structured questionnaires by interview 2-4 days postnatally. Defined as breast milk only and any breast milk. **Breast feeding duration:** measured by telephone interview at six months, postnatally using structured questionnaires. Definitions of breast feeding: breast milk only, any breast milk, and exclusive breast feeding.
Ickovics 2007	A multisite randomized controlled trial was conducted at two university-affiliated hospital prenatal clinics in Connecticut and	1,047 primarily non-employed African American pregnant women aged 14−25 years. Experimental (n=653), control (n=394). **Inclusion criteria:** less than 24 weeks	**Experimental condition:** • General antenatal education. • 10 sessions each lasting 2 hours. • Sessions from gestation week 16-40. • Approximately 8 women in each group. • Content: Group prenatal care across the pregnancy. Focus was on discussion between women and clinicians, and education and skills building to address explicit	All relevant outcomes were measured during third trimester (on average in gestation week 35). **Prenatal distress:** measured with the Pregnancy Distress Questionnaire. **Readiness for labor and delivery:** No description of measurement tool. **Readiness for infant care:** No

	Atlanta, USA.	of gestation, age 25 years or less, no medical problems requiring individualized care as "high-risk pregnancy", English or Spanish language, and willingness to be randomized.	learning objectives in prenatal care, child birth preparation, and postpartum and parenting roles as well as self-care activities on of weight and blood pressure assessment. • Led by a trained practitioner (midwife or obstetrician). **Control condition:** Individual prenatal care across the pregnancy occurs over the course of approximately 2 hours in total.	description of measurement tool. **Prenatal knowledge:** measured by a tool developed for the study to assess prenatal and infant care knowledge.
Kistin 1990	Individually randomized trial with 2 arms – 'breast feeding classes' and 'individual sessions', Conducted in Chicago, USA.	74 black women born in the US attending a midwife prenatal clinic before their 24th week of gestation. Experimental (n=38), control (n=36).	**Experimental condition:** • 50-80 minute breast-feeding class with group discussion. Participants attended at least one (more if they wished). • Topics related to breast feeding/formula use plans, health benefits of breast milk, and common challenges related to breast feeding and how to overcome them. • Sessions led by the authors. • No information on class-sizes or gestation age for education provided. **Control condition:** One-to-one contact with a medical doctor for 15 to 30 minutes before gestation week 30. The topics discussed were the same as in the experimental group.	**Breast feeding initiation:** measured at an interview in the hospital less than five days postnatally. Breast feeding defined as one or more breast feedings per day. **Breast feeding duration:** measured by self-reporting. Defined as any breast feeding for 12 weeks or longer.
Lara 2010	Individually randomized trial conducted in Mexico City, Mexico.	377 low-income pregnant women in high risk of depression. Experimental (n=250), control (n=127). **Inclusion criteria:** ≥18 years old, ≤26 weeks pregnant, having completed primary school, did not have any substance abuse, bipolar conditions or reported suicide attempts during the last 6 months, living in the metropolitan area of Mexico City, and meeting criteria for high risk for depression, based on a score of 16 or higher on the Center for Epidemiologic Studies Depression Scale (CES-D) and/or having a self-reported history of depression (only criteria in 43.2 % of the participants). **Exclusion criteria:** current	**Experimental condition:** • Psycho-educational program to prevent post-partum depression. • Eight, two-hour weekly sessions. • 5-10 participants per group. • The intervention program had several components: acknowledgement and discussion, as opposed to a formal lecture, of the "normal" perinatal period and risk factors for postpartum depression; increasing positive thinking and pleasant activities; improving self-esteem, and increasing self-care. • Four facilitators delivered the intervention - all of them had extensive clinical experience. **Control condition:** Standard care as provided by the institutions, including individual prenatal health care (checking for blood pressure, weight check etc.). In some sites they received individual talks on prenatal health care and breathing exercises to use during labor.	**Depressive symptoms:** measured by interview six weeks postnatal with the second edition of the Beck Depression Inventory (BDI-II), a 21 item self-report instrument that explores presence of symptoms during the last two weeks. A cut-off point of ≥14 was used.

		depression.		
Le 2011	Two-cited individually randomized trial conducted in Washington, DC, USA.	217 predominantly Central American immigrant women in high risk of depression. Experimental (n=112), control (n=105). **Inclusion criteria:** 18 –35 years old, ≤24 weeks gestation; no smoking, alcohol, or illicit substance abuse; and at high risk for depression, defined as scoring 16 or higher on the Center for Epidemiological Studies Depression Scale (CES–D) and/or with a self-reported personal or family history of depression. **Exclusion criteria:** current diagnosis of major depressive disorder, substance abuse, psychosis, a serious medical condition, and/or other significant psychosocial problems.	**Experimental condition:** • Psycho-educational group sessions. • Eight weekly 2-hour sessions • Content: teaching women mood regulation skills to prevent perinatal depression. • The course was taught in Spanish by one or two postbachelor's trained bilingual and/or bicultural research staff. • No information on class-sizes or gestation age for education provided. **Control condition:** Standard care as provided by their prenatal care providers at the clinic. They had the option to participate in group prenatal care classes, if they chose to in addition to their care with a midwife or MD.	**Depressive symptoms:** measured by interview six weeks postnatal with the second edition of the Beck Depression Inventory (BDI-II), a 21 item self-report instrument that explores presence of symptoms during the last two weeks.
Maimburg 2010	Individually randomized controlled trial conducted in Aarhus, Denmark.	1193 primarily middle-high education level women. Experimental (n=603), control (n=590). **Inclusion criteria:** primipara women registered at the Aarhus Midwifery Clinic in Denmark, older than 18 years of age at enrolment, singleton pregnancy, and able to speak and understand Danish.	**Experimental condition:** • General antenatal education. • Program comprised 3 modules, each lasting 3 hours. • Sessions between 30 and 35 weeks of gestation. • The form was information, video films and group discussions. • Program covered pregnancy issues, birth process, pain delivery, care for the newborn, breast feeding, the transition to parenthood, and postnatal depression. The woman's partner was also invited to participate • The instructors were midwifes of varying seniority. **Control condition:** Standard care containing individual consultations with a midwife. No offer on antenatal education but allowed to take other (private) antenatal training. The experimental condition received this offer too,	**Pain relief:** collected from the local hospital database. Overall use of pain relief, overall use of non-pharmacological pain relief, water immersion, acupuncture, intracutaneous sterile water injection, overall use of pharmacological pain relief, nitrous oxide/oxygen, intramuscular morphine, pudendal nerve block, epidural analgesia, other. **Obstetric interventions:** collected from the local hospital database. Labor induction, oxytocin augmentation, vacuum extraction, caesarean section (elective, emergency). **Sufficient knowledge about breast feeding:** measured six weeks postnatal by one question developed for the study. Measured on a 5-point Likert scale. **Breast feeding self-efficacy:** measured

				with the BSES-SF six weeks postnatally. BSES-SF is a 14-item tool measuring confidence in breast feeding answered on a 5-point Likert scale. **Breast feeding duration:** measured six weeks and six months postnatal. Breast feeding is both defined as any and exclusive breast feeding. **Postnatal depression:** measured six weeks postnatal by the self-administered 13-item questionnaire; Edinburgh Postnatal Depression Scale. A cut-off point of ≥12 was used as an indicator of risk of postnatal depression.
Maycock 2013	Individually randomized multicenter trial conducted in 8 hospitals in Perth, Australia.	1575 (863 women and 712 men) participants were recruited from 8 public hospitals. No information of number of participants randomized to experimental and control group. The final analysis was completed with 353 women in the experimental and 298 in the control condition. **Inclusion criteria:** mothers ≥18 years. Fathers had to be contactable by telephone or email, reside within Western Australia; and intend to participate in the rearing of their child.	**Experimental condition:** • One breast feeding group session for fathers. • 2-hour session with an average group size of 6 fathers. • The main topics of this session were the role of the father, the importance and benefits of breast feeding, and what to expect in the first four weeks at home with a new baby. • From birth and the following 6 weeks, the experimental group of men received written materials aimed to enhance the support for their partner's breast feeding. • Facilitated by a male instructor. **Control condition:** Standard care consisting of routine antenatal classes incorporating information on labor, birth, pain relief and breastfeeding. The experimental condition received this offer too,	**Breast feeding:** measured by questionnaire six weeks postnatal. Breast feeding defined as any breast feeding and exclusive breast feeding.
Noel-Weiss 2006	Individually randomized trial conducted in Ontario, Canada.	101 primarily middle-high SES women. 99 % lived in a supportive relationship and 87 % had made decision to breast feed prior to pregnancy. No information of number of participants randomized to experimental and control group. The final analysis was completed with 47 women in the experimental and 45 in the control condition.	**Experimental condition:** • One prenatal breast feeding workshop lasting 2.5-hour. • Participation of 2-8 women and their partners. • Conducted after gestation week 34. • The form was both based on practical breast feeding exercises with a doll, group discussions, and watching videos with breast feeding. • Facilitated by a registered nurse who had specialized in providing maternity care and breast feeding support. The facilitator was skilled with leading group discussions and providing individual counseling. **Control condition:** Standard care including the	**Maternal breast feeding self-efficacy:** measured with the BSES-SF four and eight weeks postnatal. BSES-SF is a 14-item tool measuring confidence in breast feeding answered on a 5-point Likert scale. **Breast feeding duration:** measured by asking the mother four and eight weeks postnatal whether she was breast feeding and how much. Breast feeding defined as exclusive breast milk and any breast milk.

		Inclusion criteria: primipara women expecting a single child, an uncomplicated birth, and planning to breastfeed, able to read and write in English and have a telephone to complete the postpartum questionnaires. **Exclusion criteria:** mother and her infant not discharged at the same time; mother not able to breastfeed without restriction.	choice of physician or midwife, frequency of prenatal visits, and attendance at prenatal classes, was defined by each mother. The experimental condition received this offer too.	
Ortiz Collado 2014	A multicentre randomized, longitudinal clinical study conducted in three cities in Spain and France.	184 primarily low SES women in risk of postnatal depression. Experimental (n=92), control (n=92). **Inclusion criteria:** middle or low socio-economic status, <20 weeks of gestation, a moderate to high risk of postnatal depression, no more than two children, no organic serious physical pathology, no psychiatric diagnosis, no alcohol or illicit substance abuse, and able to understand the language. **Exclusion criteria:** having a current diagnosis of psychiatric disorder or a serious medical condition.	**Experimental condition:** • Preparation for parenthood group sessions. • 6-8 couples met for 10 weekly sessions each lasting 2 hours and 15 minutes. • Began during the second term of pregnancy. • The classes involved work on individual feelings and affective bonds, with specific objectives for the man and the woman in each participating couple. The program was focusing on preparation for parenting and not just for the childbirth, as well as preparation for both the mother and the father. Each session consisted of an interactive exchange of information (60%) and practical exercises (40%). • No information on educators. **Control condition:** Standard antenatal education program consisting of eight sessions of two hours each during the third term of pregnancy. The focus was childbirth and pregnancy health. Each group was open and could receive 12 couples or more.	All outcomes were measured by self-reported questionnaires mailed between five and 12 weeks postnatally. **Depressive symptoms:** Measured by the self-administered 10-item questionnaire; Edinburgh Postnatal Depression Scale. **Amount of social support received:** Measured by the 11-item Functional Social Support Questionnaire. The questionnaire refers to two dimensions of functional social support: affective support and confidant support. Satisfaction with different situations is rated on a 5-point Likert scale. **Relationship with partner**: Measured by the 32-item Dyadic Adjustment Scale (DAS). Both mothers and fathers rated items on various Likert-type scales.
Rossiter 1994	Individually randomized trial conducted in Sydney, Australia.	194 Vietnamese, primarily unemployed, low SES pregnant women. Experimental (n=108), control (n=86). **Inclusion criteria:** ethnic Vietnamese or other women who were born and reared in Vietnam; Vietnamese speaking; at least 12 weeks pregnant; gave consent to participate. **Exclusion criteria:** unforeseen	**Experimental condition:** • Breast feeding education program. 3 sessions each lasting 2 hours. • Content: a 25-minute videotape followed by small-group discussion sessions. The aims were to provide information on the benefits of breast feeding, relate this information to the women's background, and discuss any misconceptions about the superiority of formula milk and the norm of infant feeding practices in Australia. • The program was conducted in Vietnamese by the parenthood educators of the hospitals, with the	**Breast feeding initiation:** measured at visit at the hospital/home visit within one week postnatally. Breast feeding defined as being the main source of nutrition. **Breast feeding duration:** assessed at home visit six months postnatally. Breast feeding defined as being the main source of nutrition.

		circumstances (miscarriage, stillbirth, change of address).	assistance of Vietnamese health interpreter. • No information on class-sizes or gestation age for education provided. **Control condition:** Participants were provided with official breast feeding and childbirth pamphlets.	
Rouhe 2013	Individually randomized trial conducted in Helsinki, Finland.	371 women with severe fear of childbirth. Experimental (n=131), control (n=240). **Inclusion criteria:** Fear of childbirth, defined as a sum score ≥100 on the Wijma Delivery Expectancy Questionnaire; primipara. **Exclusion criteria:** manifest psychosis; severe depression; serious problems of alcohol or drug abuse.	**Experimental condition:** • Group psycho-educational classes to reduce fear of birth. • Six 2-hour group sessions during pregnancy from 26th to 35th week gestation and one session 6–8 weeks after delivery. • The focus of the intervention was on increasing individual independence and awareness of one's own abilities, the choices available during one's delivery and the successful transition to motherhood. Partners participated in one of the group sessions. • Group size: maximum six women, • Instructor: a psychologist. • No planned visits with an obstetrician. **Control condition:** A letter in which they were recommended to discuss their fear of childbirth in their primary healthcare maternity unit. Referral to a fear of childbirth team.	All outcomes related to delivery were derived from hospital obstetric patient records. **Pain relief:** epidural analgesia. **Obstetric interventions:** Spontaneous vaginal delivery, induction of labor, caesarean section (elective, emergency), vacuum extraction.
Schulz 2006	Individually randomized trial conducted in California, USA.	52 primarily European American couples. Experimental (n=28 couples), control (n=24 couples). **Inclusion criteria:** couples living together, expecting their first child, and over 18 years of age.	**Experimental condition**: • Couple-focused intervention for partners becoming parents. • 24 weekly 2.5 hour couple group sessions. • Sessions from 3 months before birth - 3 months after birth. • Each group included 4 couples and one co-leader married couple. • Topics for discussion included: how participants viewed themselves and their relationships, division of family labor, communication and problem-solving styles, and relationship as a couple. **Control condition:** Standard care, including home and lab visits (included interviews, interaction tasks, and cognitive assessments of the children after birth), The experimental condition received this offer too.	**Marital satisfaction:** measured by the MAT questionnaire containing 16 items. Both men and women answered this questionnaire at six months and 66 months postnatally. **Divorce/separation:** reported at 66 months postnatally.
Werner 2013	Individually randomized single-blind controlled trial	727 women Experimental (n=497), control (n = 230).	**Experimental condition:** • Self-hypnosis for childbirth-education. • Three 1-hour sessions held over three consecutive weeks. The first session lasted 2.5 hours.	All outcomes related to delivery were extracted from an ongoing data collection from all births at the hospital or from medical records.

	with 3 arms: one experimental group, an active comparison group and one control group. In this review effects of the most intensive group-based program (hypnosis) against the control group is tested. Conducted in Aarhus, Denmark.	**Inclusion criteria:** no chronic diseases, uncomplicated pregnancy, primipara, older than 18 years, and able to understand and speak Danish.	• The program included three audiorecordings including a 20-minute section especially meant for labor. • Classes were taught by midwifes trained in hypnosis. • No information on group size provided. **Control condition:** Standard care consisting of 4-5 individual consultations with a midwife and a tour of the birth department. The experimental condition received this offer too.	**Pain relief:** use of epidural analgesia during birth. **Obstetric interventions:** spontaneous vaginal birth, cesarean section (elective, emergency), oxytocin augmentation, vacuum extraction. **Breast feeding duration (any breast feeding):** derived from questionnaires four months postnatally.
Westney 1988	Individually randomized trial.	A volunteer sample of 28 black, unmarried, 15-18 year old prospective fathers. Experimental (n=15), control (n=13).	**Experimental condition:** • Prenatal classes 4 times weekly each lasting 2 hours. • Classes addressed human sexuality, pregnancy and prenatal care, labor and delivery, infant and child care. • Teaching approaches included lectures, audiovisual aids, and group discussions of concerns. • Presented by a female registered nurse-specialist in maternal-child care. • No information on class-sizes or gestation age for education provided. **Control condition:** Standard care (content not described). No participants reported to participate in any other pregnancy-related education program.	**Paternal knowledge of human sexuality, pregnancy and prenatal care, labor and delivery, infant and child care, and support towards the mother:** measured after last experimental session (gestation week unknown) using a 75-item questionnaire developed for the study.

Effect tables

Comparison 1: Depression prevention classes versus standard care

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
1.1 Depression 3 months postnatal (EPDS)	1	190	Risk Ratio (M-H, Fixed, 95% CI)	0.85 [0.46, 1.59]
1.2 Depression 3 months postnatal (GHQ-D tool)	1	190	Risk Ratio (M-H, Fixed, 95% CI)	1.17 [0.70, 1.95]
1.3 Depression 3 months postnatal (Scan tool)	1	190	Risk Ratio (M-H, Fixed, 95% CI)	0.51 [0.13, 1.98]
1.4 Depressive symptoms in pregnancy (BDI-II tool)	1	186	Mean Difference (IV, Fixed, 95% CI)	-2.10 [-4.61, 0.41]
1.5 Depressive symptoms 6 weeks postnatal (BDI-II tool)	1	180	Mean Difference (IV, Fixed, 95% CI)	0.31 [-2.10, 2.72]
1.6 Depressive symptoms 6 weeks postnatal (> 14 on BDI-II tool)	1	149	Risk Ratio (M-H, Fixed, 95% CI)	0.95 [0.62, 1.45]
1.7 High vs. low confidence in ability to solve problems 3 months postnatal	1	190	Risk Ratio (M-H, Fixed, 95% CI)	0.71 [0.28, 1.80]
1.8 High vs. low belief in personal control when solving problems 3 months postnatal	1	190	Risk Ratio (M-H, Fixed, 95% CI)	1.07 [0.80, 1.43]
1.9 High vs. low belief in internal factors influencing their life 3 months postnatal	1	190	Risk Ratio (M-H, Fixed, 95% CI)	0.98 [0.73, 1.31]

Comparison 2: Psycho-social prevention program versus brochure

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
2.1 Depressive symptoms women 6 months postnatal (CESDS)	1	152	Mean Difference (IV, Fixed, 95% CI)	-0.08 [-0.21, 0.05]
2.2 Depressive symptoms men 6 months postnatal (CESDS)	1	152	Mean Difference (IV, Fixed, 95% CI)	0.02 [-0.07, 0.11]
2.3 Co-parental support women 6 months postnatal	1	152	Mean Difference (IV, Fixed, 95% CI)	0.30 [-0.04, 0.64]
2.4 Co-parental support men 6 months postnatal	1	152	Mean Difference (IV, Fixed, 95% CI)	0.29 [0.05, 0.53]
2.5 Co-parental undermining women 6 months postnatal	1	152	Mean Difference (IV, Fixed, 95% CI)	0.04 [-0.20, 0.28]
2.6 Co-parental undermining men 6 months postnatal	1	152	Mean Difference (IV, Fixed, 95% CI)	0.10 [-0.16, 0.36]
2.7 Parenting-based closeness women 6 months postnatal	1	152	Mean Difference (IV, Fixed, 95% CI)	0.06 [-0.30, 0.42]
2.8 Parenting-based closeness men 6 months postnatal	1	152	Mean Difference (IV, Fixed, 95% CI)	0.35 [0.04, 0.66]
2.9 Anxiety women 6 months postnatal	1	152	Mean Difference (IV, Fixed, 95% CI)	-0.11 [-1.53, 1.31]
2.10 Anxiety men 6 months postnatal	1	152	Mean Difference (IV, Fixed, 95% CI)	-0.79 [-2.02, 0.44]
2.11 Mother-child dysfunctional	1	152	Mean Difference (IV, Fixed, 95%	-0.10 [-0.25, 0.05]

interaction 6 months postnatal			CI)	
2.12 Father-child dysfunctional interaction 6 months postnatal	1	152	Mean Difference (IV, Fixed, 95% CI)	-0.26 [-0.43, -0.09]

Comparison 3: Psycho-educational classes versus letter

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
3.1 Epidural analgesia	1	371	Risk Ratio (M-H, Fixed, 95% CI)	1.13 [0.96, 1.33]
3.2 Spontaneous delivery	1	371	Risk Ratio (M-H, Fixed, 95% CI)	1.33 [1.11, 1.61]
3.3 Overall caesarean section	1	371	Risk Ratio (M-H, Fixed, 95% CI)	0.70 [0.49, 1.01]
3.4 Elective caesarean section	1	371	Risk Ratio (M-H, Fixed, 95% CI)	0.83 [0.46, 1.50]
3.5 Emergency caesarean section	1	371	Risk Ratio (M-H, Fixed, 95% CI)	0.62 [0.37, 1.06]
3.6 Vacuum extraction	1	371	Risk Ratio (M-H, Fixed, 95% CI)	0.69 [0.42, 1.13]
3.7 Induction of labor	1	371	Risk Ratio (M-H, Fixed, 95% CI)	1.10 [0.74, 1.64]

Comparison 4: Program with psycho-somatic approach versus standard antenatal education program

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
4.1 Depressive symptoms 5-12 weeks postnatal (EPDS)	1	127	Mean Difference (IV, Fixed, 95% CI)	-1.77 [-3.75, 0.21]
4.2 Lack of social support 5-12 weeks postnatal(Functional Social Support Questionnaire)	1	127	Mean Difference (IV, Fixed, 95% CI)	-1.61 [-4.66, 1.44]
4.3 Dissatisfaction with relationship women 5-12 weeks postnatal (DASS)	1	127	Mean Difference (IV, Fixed, 95% CI)	5.38 [-4.07, 14.83]
4.4 Dissatisfaction with relationship men 5-12 weeks postnatal (DASS)	1	127	Mean Difference (IV, Fixed, 95% CI)	4.30 [-1.21, 9.81]

Comparison 5: Couple-focused classes versus standard care

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
5.1 Marital satisfaction women 6 months postnatal	1	49	Mean Difference (IV, Fixed, 95% CI)	-5.54 [-16.95, 5.87]
5.2 Marital satisfaction men 6 months postnatal	1	49	Mean Difference (IV, Fixed, 95% CI)	-0.50 [-9.77, 8.77]
5.3 Marital satisfaction women 5.5 years postnatal	1	29	Mean Difference (IV, Fixed, 95% CI)	8.90 [-10.47, 28.27]
5.4 Marital satisfaction men 5.5 years postnatal	1	29	Mean Difference (IV, Fixed, 95% CI)	5.33 [-9.58, 20.24]
5.5 Divorce/separation 5.5 years postnatal	1	45	Risk Ratio (M-H, Fixed, 95% CI)	1.20 [0.37, 3.88]

Comparison 6: Self-hypnosis classes versus standard care

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
6.1 Epidural analgesia	1	723	Risk Ratio (M-H, Fixed, 95% CI)	1.04 [0.82, 1.32]
6.2 Spontaneous delivery	1	723	Risk Ratio (M-H, Fixed, 95% CI)	1.00 [0.90, 1.11]
6.3 Overall caesarean section	1	723	Risk Ratio (M-H, Fixed, 95% CI)	1.25 [0.89, 1.76]
6.4 Elective caesarean section	1	723	Risk Ratio (M-H, Fixed, 95% CI)	0.51 [0.22, 1.19]
6.5 Emergency caesarean section	1	723	Risk Ratio (M-H, Fixed, 95% CI)	1.52 [1.02, 2.27]
6.6 Vacuum extraction	1	723	Risk Ratio (M-H, Fixed, 95% CI)	0.75 [0.51, 1.10]
6.7 Oxytocin augmentation	1	723	Risk Ratio (M-H, Fixed, 95% CI)	1.09 [0.90, 1.30]
6.8 Labor induction	1	723	Risk Ratio (M-H, Fixed, 95% CI)	0.91 [0.79, 1.04]
6.9 Any breast feeding 4 months postnatal	1	698	Risk Ratio (M-H, Fixed, 95% CI)	0.97 [0.91, 1.04]

Comparison 7: General antenatal education classes versus standard care

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
7.1 Overall pain relief	1	1162	Risk Ratio (M-H, Fixed, 95% CI)	1.00 [0.94, 1.05]
7.2 Overall pharmacological pain relief	1	1162	Risk Ratio (M-H, Fixed, 95% CI)	0.93 [0.86, 1.01]
7.3 Epidural analgesia	1	1162	Risk Ratio (M-H, Fixed, 95% CI)	0.84 [0.73, 0.98]
7.4 Nitrous oxide/oxygen	1	1162	Risk Ratio (M-H, Fixed, 95% CI)	1.03 [0.90, 1.19]
7.5 Intramuscular morphine	1	1162	Risk Ratio (M-H, Fixed, 95% CI)	0.68 [0.29, 1.57]
7.6 Pudendal nerve block	1	1162	Risk Ratio (M-H, Fixed, 95% CI)	1.27 [0.78, 2.07]
7.7 Other pharmacological (primary halcion, codein, paracetamol)	1	1162	Risk Ratio (M-H, Fixed, 95% CI)	0.89 [0.63, 1.24]
7.8 Overall non-pharmacological pain relief	1	1162	Risk Ratio (M-H, Fixed, 95% CI)	1.04 [0.94, 1.15]
7.9 Water immersion	1	1162	Risk Ratio (M-H, Fixed, 95% CI)	1.00 [0.87, 1.16]
7.10 Acupuncture	1	1162	Risk Ratio (M-H, Fixed, 95% CI)	0.94 [0.79, 1.14]
7.11 Intracutaneous sterile water injection	1	1162	Risk Ratio (M-H, Fixed, 95% CI)	0.88 [0.67, 1.16]
7.12 Spontaneous delivery	1	1162	Risk Ratio (M-H, Fixed, 95% CI)	1.03 [0.94, 1.13]
7.13 Overall caesarean section	1	1162	Risk Ratio (M-H, Fixed, 95% CI)	0.91 [0.73, 1.15]
7.14 Elective caesarean section	1	1162	Risk Ratio (M-H, Fixed, 95% CI)	0.98 [0.57, 1.68]
7.15 Emergency caesarean section	1	1162	Risk Ratio (M-H, Fixed, 95% CI)	0.90 [0.69, 1.17]
7.16 Vacuum extraction	1	1162	Risk Ratio (M-H, Fixed, 95% CI)	1.03 [0.80, 1.33]
7.17 Oxytocin augmentation	1	1162	Risk Ratio (M-H, Fixed, 95% CI)	0.97 [0.87, 1.08]
7.18 Labor induction	1	1162	Risk Ratio (M-H, Fixed, 95% CI)	1.02 [0.91, 1.15]
7.19 Sufficient knowledge about breast feeding	1	1060	Risk Ratio (M-H, Fixed, 95% CI)	1.08 [1.01, 1.15]
7.20 Exclusive breast feeding 6 weeks postnatal	1	1048	Risk Ratio (M-H, Fixed, 95% CI)	1.01 [0.95, 1.07]
7.21 Any breast feeding 6 weeks postnatal	1	836	Risk Ratio (M-H, Fixed, 95% CI)	1.01 [0.98, 1.04]
7.22 Exclusive breast feeding 6 months postnatal	1	1048	Risk Ratio (M-H, Fixed, 95% CI)	0.88 [0.55, 1.41]
7.23 Any breast feeding 6 months postnatal	1	836	Risk Ratio (M-H, Fixed, 95% CI)	1.02 [0.92, 1.12]
7.24 Postnatal depression 6 weeks postnatal (EPDS)	1	1069	Risk Ratio (M-H, Fixed, 95% CI)	0.90 [0.59, 1.37]

7.25 Breast feeding self-efficacy 6 weeks postnatal	1	1058	Mean Difference (IV, Fixed, 95% CI)	0.03 [-0.05, 0.11]

Comparison 8: Group prenatal care (20 hours) versus individual prenatal care (2 hours)

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
8.1 Prenatal and infant care knowledge	1	934	Mean Difference (IV, Fixed, 95% CI)	2.60 [1.68, 3.52]
8.2 Readiness for labor and delivery	1	934	Mean Difference (IV, Fixed, 95% CI)	7.60 [3.34, 11.86]
8.3 Readiness for infant care	1	934	Mean Difference (IV, Fixed, 95% CI)	3.10 [-0.14, 6.34]
8.4 Prenatal distress	1	934	Mean Difference (IV, Fixed, 95% CI)	-0.40 [-1.33, 0.53]

Comparison 9: Paternal education class versus standard care

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
9.1 Paternal knowledge	1	28	Mean Difference (IV, Fixed, 95% CI)	9.55 [1.25, 17.85]
9.2 Exclusive breast feeding 6 weeks postnatal	1	651	Risk Ratio (M-H, Fixed, 95% CI)	1.04 [0.88, 1.23]
9.3 Any breast feeding 6 weeks postnatal	1	651	Risk Ratio (M-H, Fixed, 95% CI)	1.09 [1.00, 1.18]

Comparison 10: Extra breast feeding sessions versus standard care

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
10.1 Spontaneous delivery	1	70	Risk Ratio (M-H, Fixed, 95% CI)	1.04 [0.77, 1.42]
10.2 Overall caesarean section	1	70	Risk Ratio (M-H, Fixed, 95% CI)	0.67 [0.12, 3.75]
10.3 Vacuum extraction	1	70	Risk Ratio (M-H, Fixed, 95% CI)	1.00 [0.36, 2.80]
10.4 Forceps	1	70	Risk Ratio (M-H, Fixed, 95% CI)	1.00 [0.15, 6.71]
10.5 Breast feeding initiation – breast milk only	1	618	Risk Ratio (M-H, Fixed, 95% CI)	0.99 [0.91, 1.08]
10.6 Breast feeding initiation – any breast milk	1	618	Risk Ratio (M-H, Fixed, 95% CI)	0.99 [0.95, 1.02]
10.7 Exclusive breast feeding 6 weeks postnatal	1	70	Risk Ratio (M-H, Fixed, 95% CI)	3.20 [1.88, 5.46]
10.8 Breast milk only 8 weeks postnatal	1	92	Risk Ratio (M-H, Fixed, 95% CI)	1.12 [0.85, 1.49]
10.9 Any breast milk 8 weeks postnatal	1	92	Risk Ratio (M-H, Fixed, 95% CI)	0.96 [0.33, 2.75]
10.10 Exclusive breast feeding 6 months postnatal	1	592	Risk Ratio (M-H, Fixed, 95% CI)	1.16 [0.67, 2.01]
10.11 Breast milk only 6 months postnatal	1	592	Risk Ratio (M-H, Fixed, 95% CI)	0.96 [0.77, 1.20]
10.12 Any breast milk 6 months postnatal	1	592	Risk Ratio (M-H, Fixed, 95% CI)	0.92 [0.79, 1.07]

10.13 Breast feeding self-efficacy 4 weeks postnatal	1	80	Mean Difference (IV, Fixed, 95% CI)	4.60 [0.72, 8.48]
10.14 Breast feeding self-efficacy 8 weeks postnatal	1	74	Mean Difference (IV, Fixed, 95% CI)	2.79 [-0.76, 6.34]

Comparison 11: Breast feeding classes versus one-to-one contact

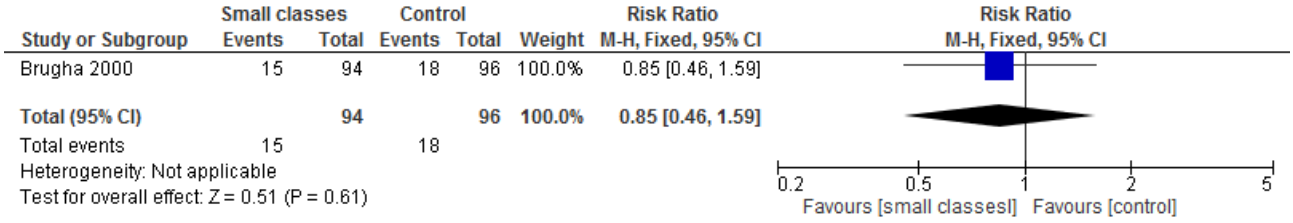
Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
11.1 Breast feeding initiation - one or more breast feedings per day	1	74	Risk Ratio (M-H, Fixed, 95% CI)	0.89 [0.55, 1.45]
11.2 Any breast feeding 12 weeks postnatal	1	74	Risk Ratio (M-H, Fixed, 95% CI)	2.84 [0.61, 13.18]

Comparison 12: Breast feeding classes versus breast feeding and childbirth pamphlets

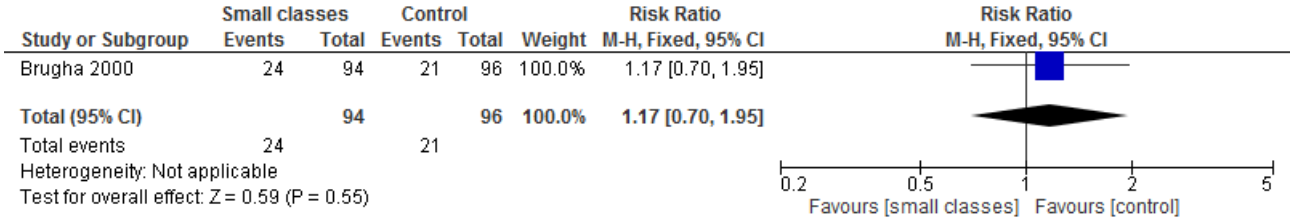
Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
12.1 Breast feeding initiation - main source of nutrition	1	178	Risk Ratio (M-H, Fixed, 95% CI)	1.86 [1.35, 2.55]
12.2 Breast feeding as main source of nutrition 6 months postnatal	1	175	Risk Ratio (M-H, Fixed, 95% CI)	1.59 [0.86, 2.94]

Forest plots

Comparison 1: Depression prevention versus standard care, outcome 1: Depression 3 months postnatal (EPDS tool)

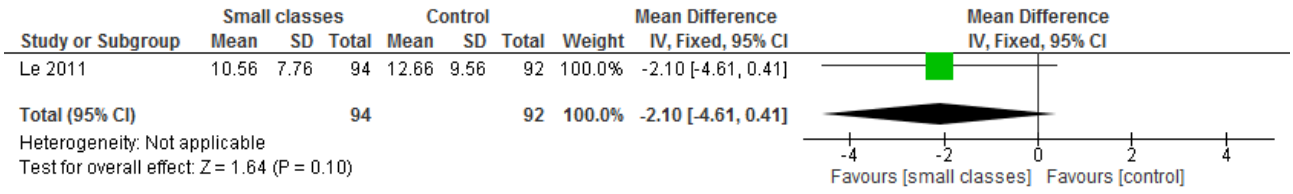


Comparison 1: Depression prevention versus standard care, outcome 2: Depression 3 months postnatal (GHQ-D tool)

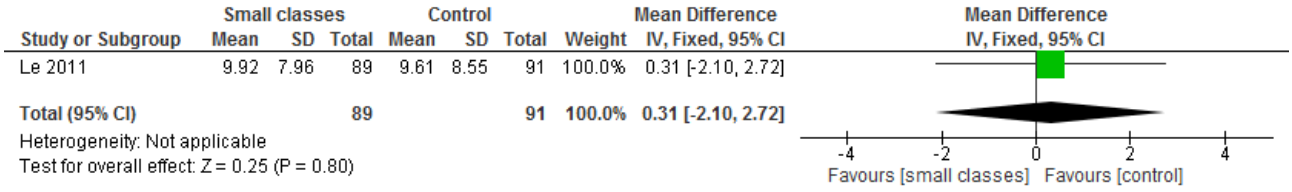


Comparison 1: Depression prevention versus standard care, outcome 3: Depression 3 months postnatal (SCAN tool)

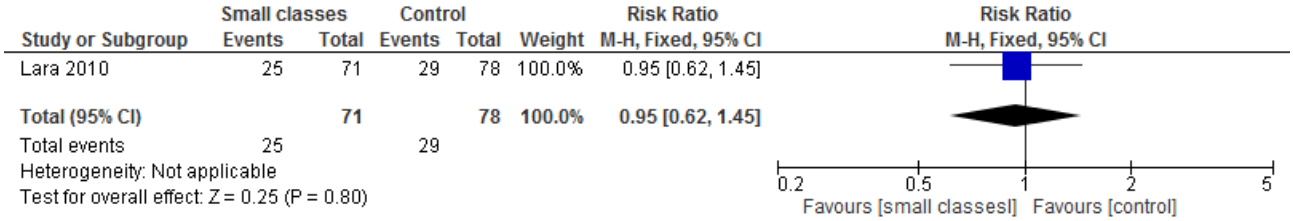
Comparison 1: Depression prevention versus standard care, outcome 4: Depressive symptoms in pregnancy (BDI-II tool)



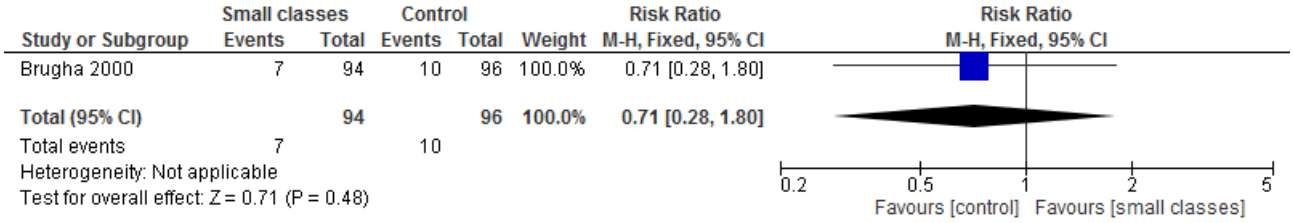
Comparison 1: Depression prevention versus standard care, outcome 5: Depressive symptoms 6 weeks postnatal (BDI-II tool)



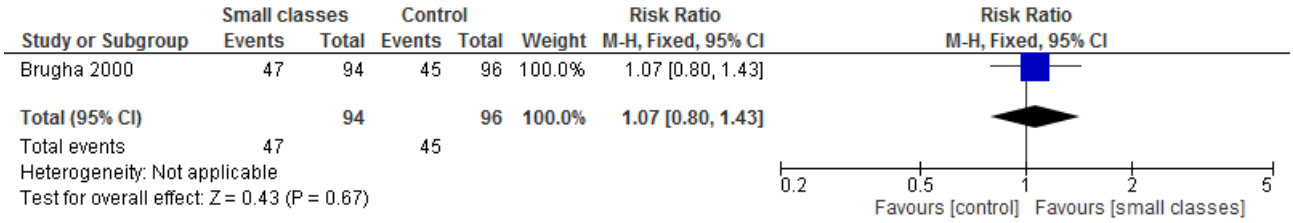
Comparison 1: Depression prevention versus standard care, outcome 6: Depressive symptoms 6 weeks postnatal (> 14 on BDI-II tool)



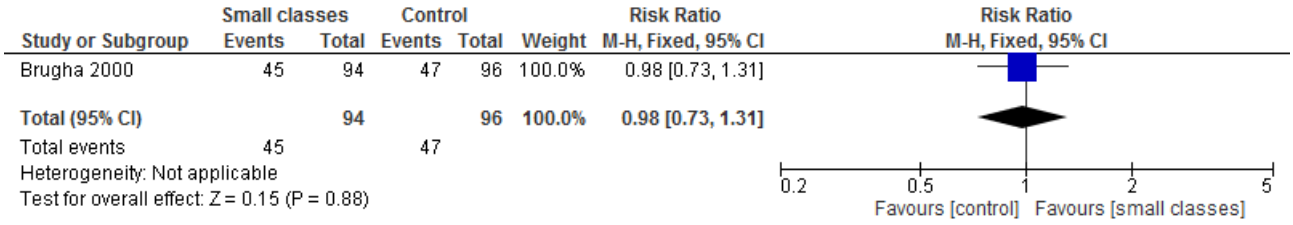
Comparison 1: Depression prevention versus standard care, outcome 7: High vs. low confidence in ability to solve problems 3 months postnatal



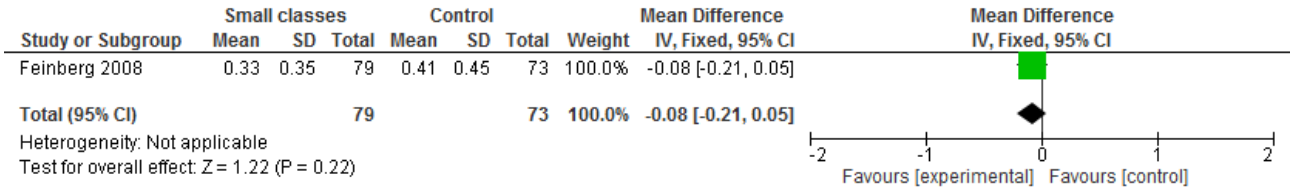
Comparison 1: Depression prevention versus standard care, outcome 8: High vs. low belief in personal control when solving problems 3 months postnatal



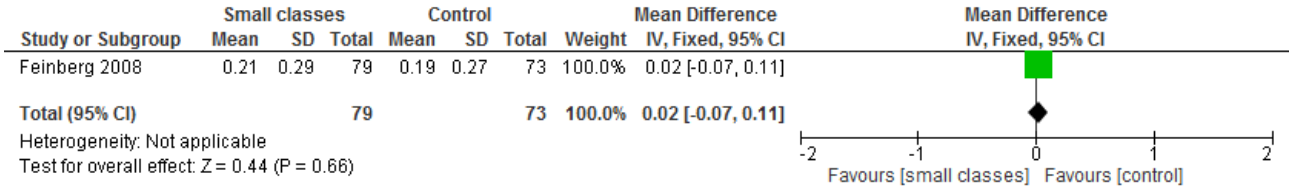
Comparison 1: Depression prevention versus standard care, outcome 9: High vs. low belief in internal factors influencing their life 3 months postnatal



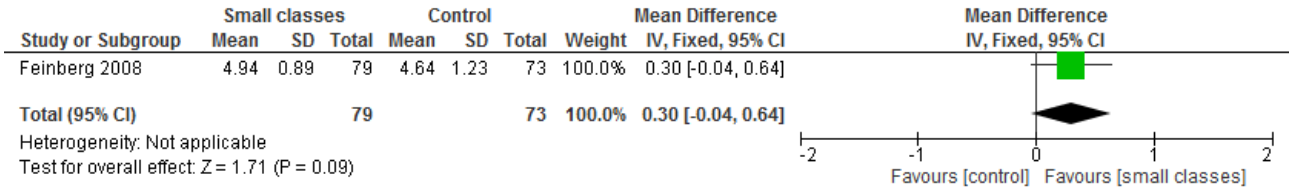
Comparison 2: Psycho-social prevention program versus brochure on child care, outcome 1: Depressive symptoms women 6 months postnatal (CESDS)



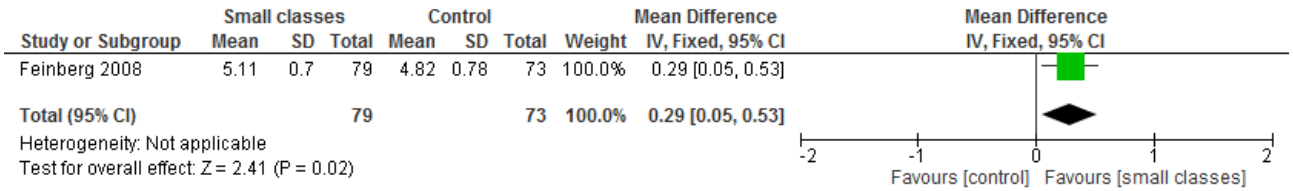
Comparison 2: Psycho-social prevention program versus brochure on child care, outcome 2: Depressive symptoms men 6 months postnatal (CESDS)



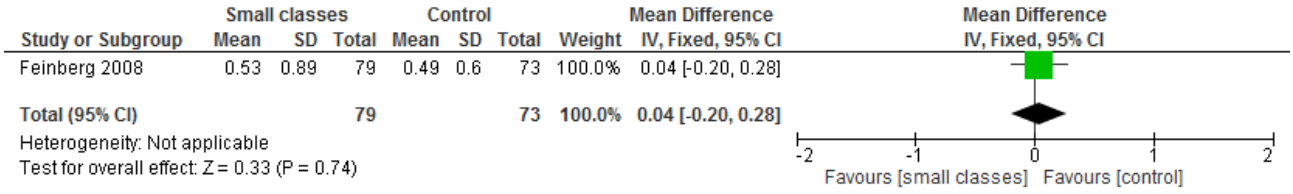
Comparison 2: Psycho-social prevention program versus brochure on child care, outcome 3: Co-parental support women 6 months postnatal



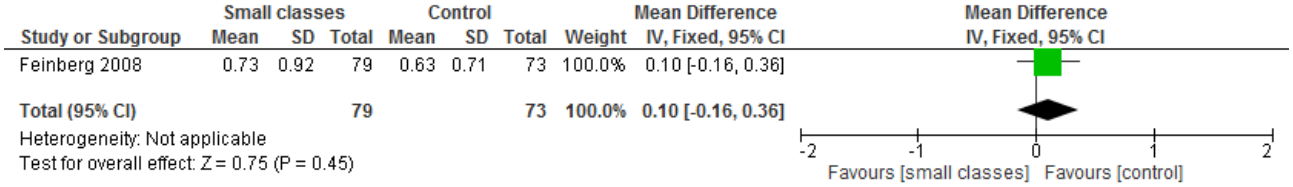
Comparison 2: Psycho-social prevention program versus brochure on child care, outcome 4: Co-parental support men 6 months postnatal



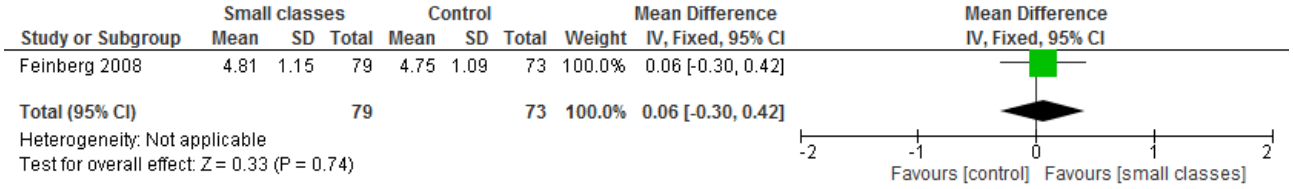
Comparison 2: Psycho-social prevention program versus brochure on child care, outcome 5: Co-parental undermining women 6 months postnatal



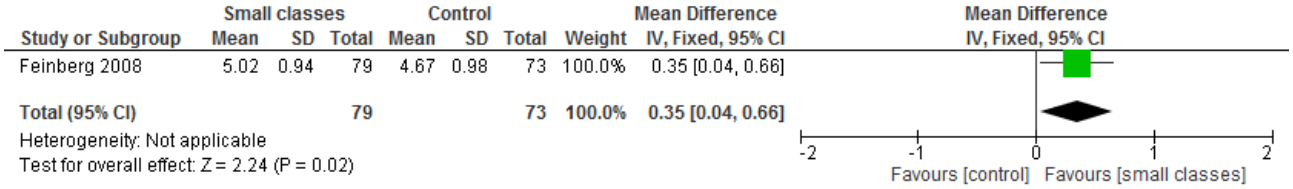
Comparison 2: Psycho-social prevention program versus brochure on child care, outcome 6: Co-parental undermining men 6 months postnatal



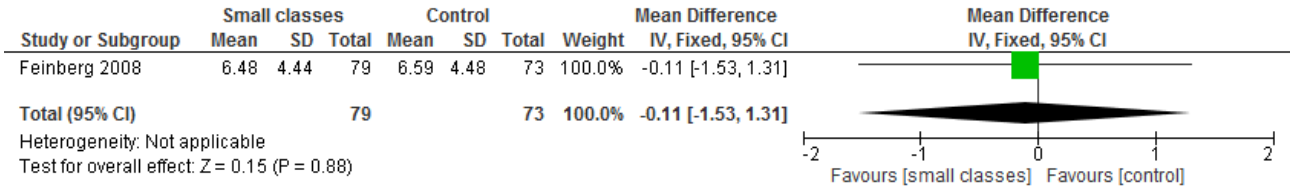
Comparison 2: Psycho-social prevention program versus brochure on child care, outcome 7: Parenting-based closeness women 6 months postnatal



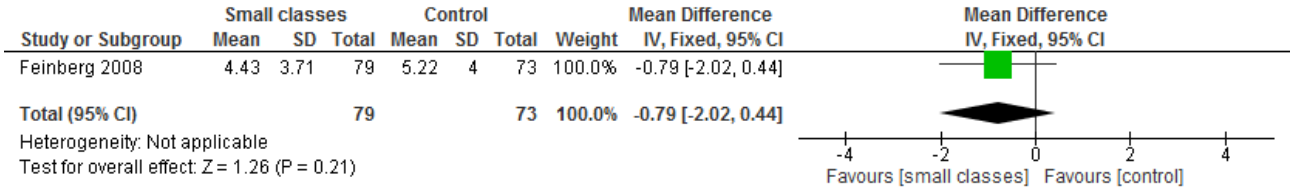
Comparison 2: Psycho-social prevention program versus brochure on child care, outcome 8: Parenting-based closeness men 6 months postnatal



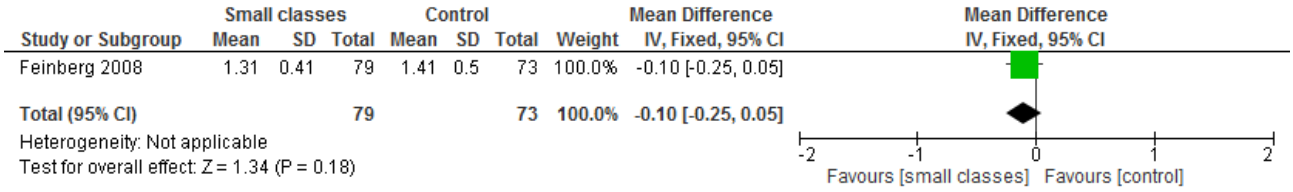
Comparison 2: Psycho-social prevention program versus brochure on child care, outcome 9: Anxiety women 6 months postnatal



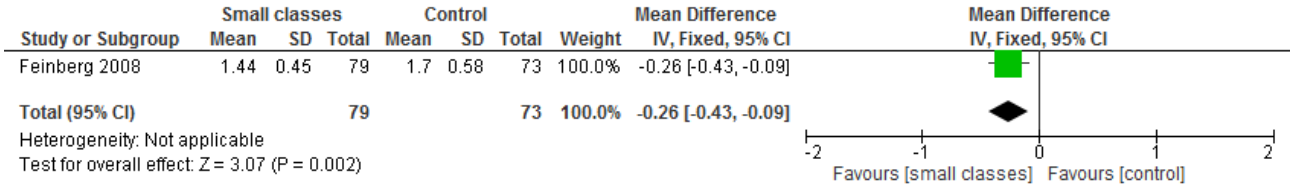
Comparison 2: Psycho-social prevention program versus brochure on child care, outcome 10: Anxiety men 6 months postnatal



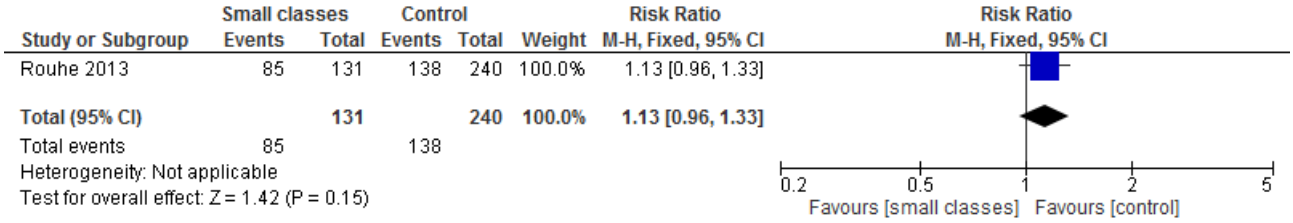
Comparison 2: Psycho-social prevention program versus brochure on child care, outcome 11: Mother-child dysfunctional interaction 6 months postnatal



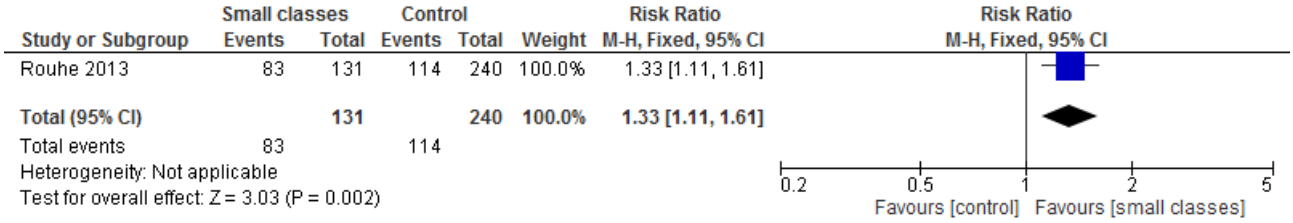
Comparison 2: Psycho-social prevention program versus brochure on child care, outcome 12: Father-child dysfunctional interaction 6 months postnatal



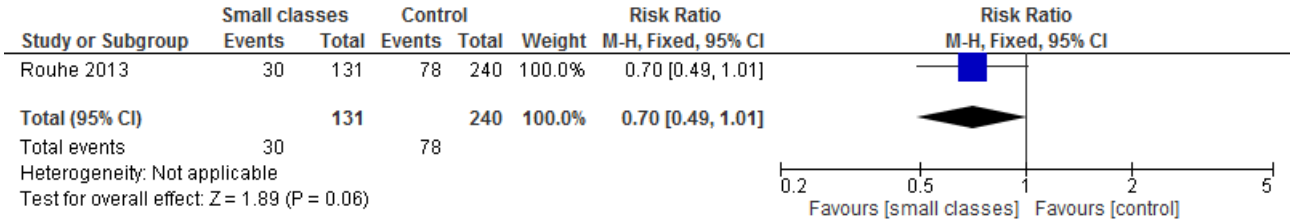
Comparison 3: Psycho-educational classes versus letter on fear of childbirth, outcome 1: Epidural analgesia



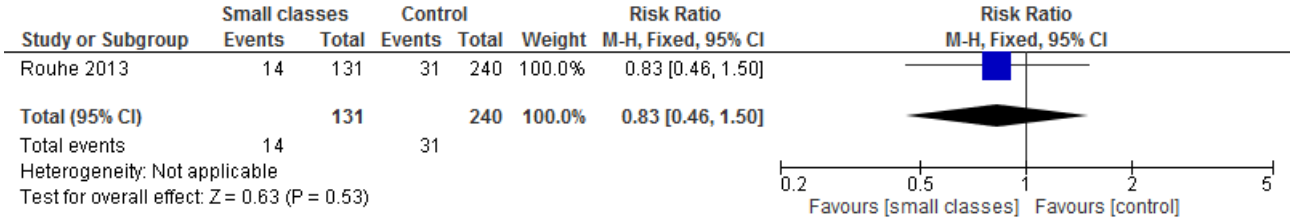
Comparison 3: Psycho-educational classes versus letter on fear of childbirth, outcome 2: Spontaneous delivery



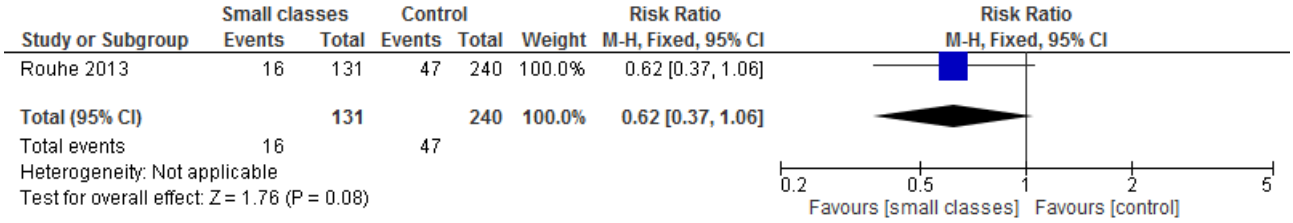
Comparison 3: Psycho-educational classes versus letter on fear of childbirth, outcome 3: Overall caesarean section



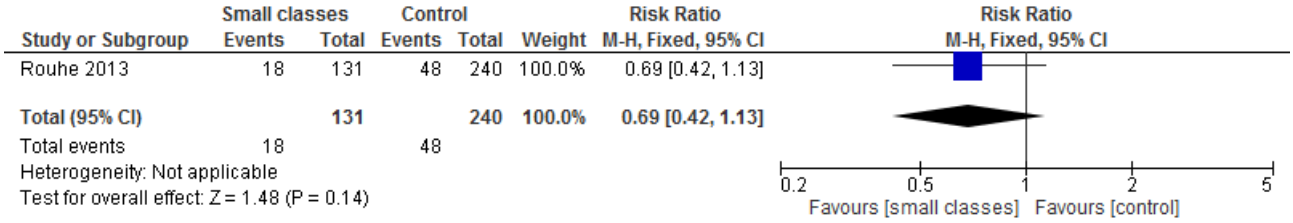
Comparison 3: Psycho-educational classes versus letter on fear of childbirth, outcome 4: Elective caesarean section



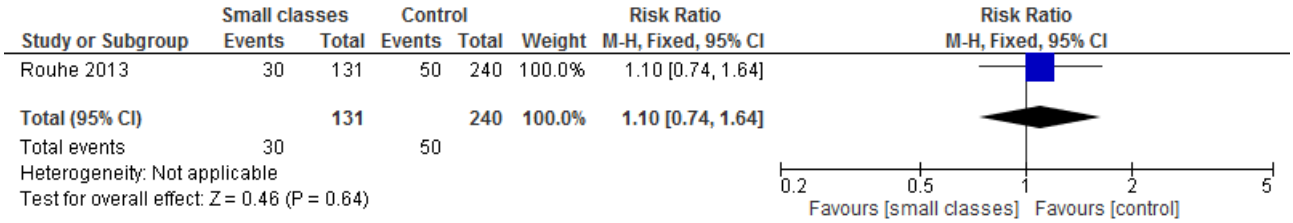
Comparison 3: Psycho-educational classes versus letter on fear of childbirth, outcome 5: Emergency caesarean section



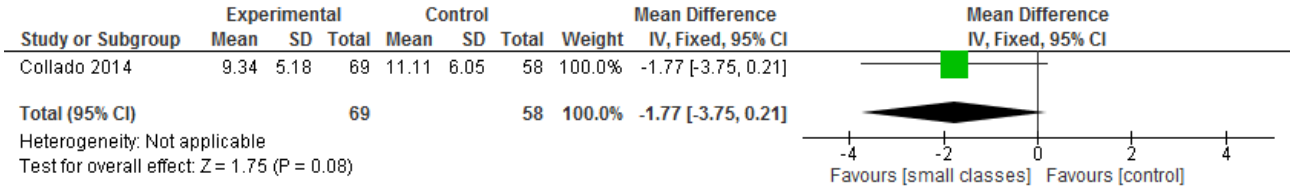
Comparison 3: Psycho-educational classes versus letter on fear of childbirth, outcome 6: Vacuum extraction



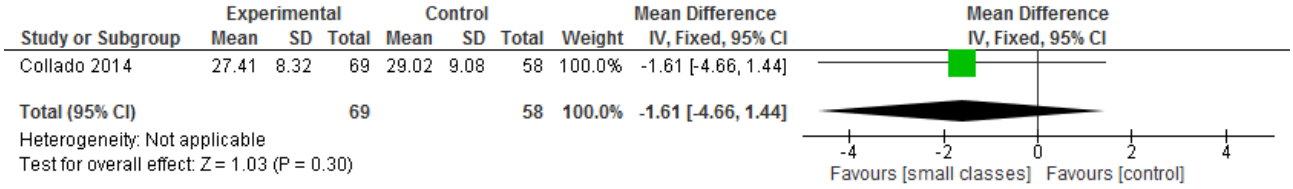
Comparison 3: Psycho-educational classes versus letter on fear of childbirth, outcome 7: Induction of labor



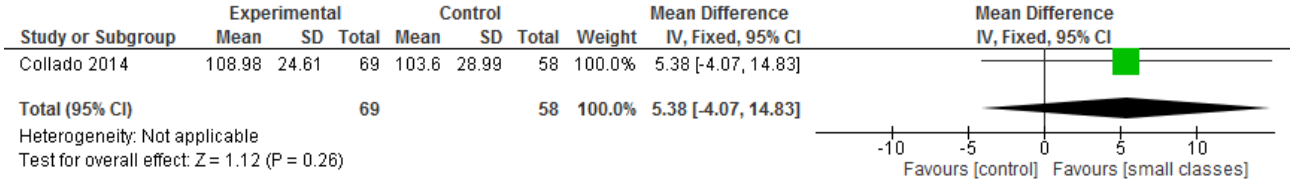
Comparison 4: Program with psycho-somatic approach versus standard antenatal education program, outcome 1: Depressive symptoms 5-12 weeks postnatal (EPDS)



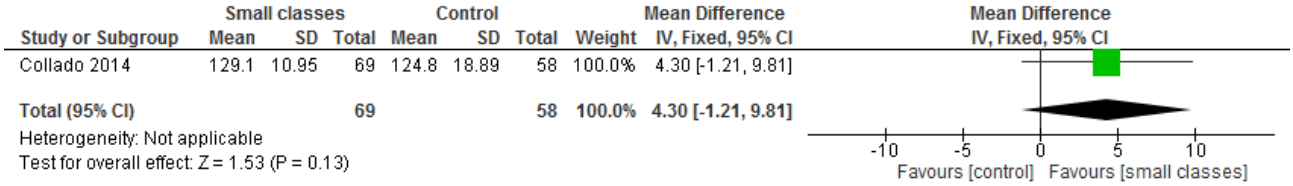
Comparison 4: Program with psycho-somatic approach versus standard antenatal education program, outcome 2: Lack of social support 5-12 weeks postnatal (Functional Social Support Questionnaire)



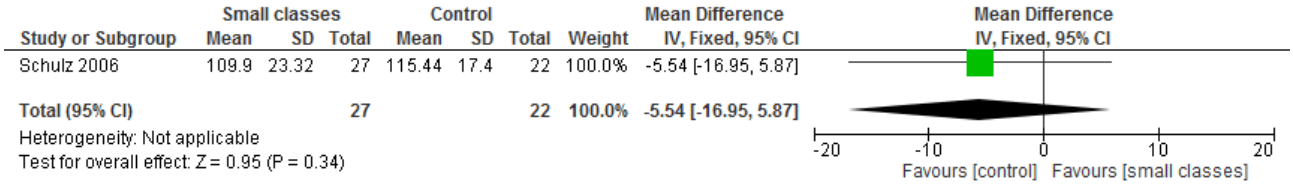
Comparison 4: Program with psycho-somatic approach versus standard antenatal education program, outcome 3: Dissatisfaction with relationship women 5-12 weeks postnatal (DASS)



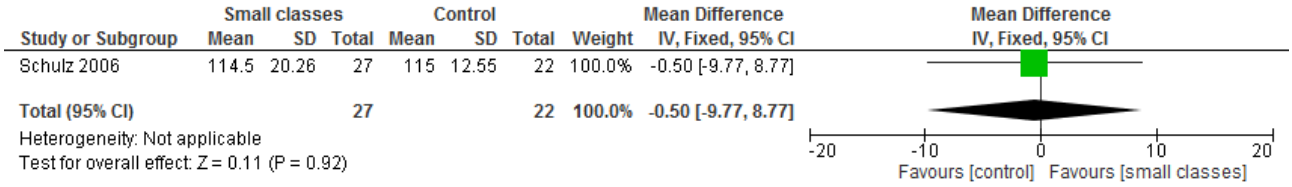
Comparison 4: Program with psycho-somatic approach versus standard antenatal education program, outcome 4: Dissatisfaction with relationship men 5-12 weeks postnatal (DASS)



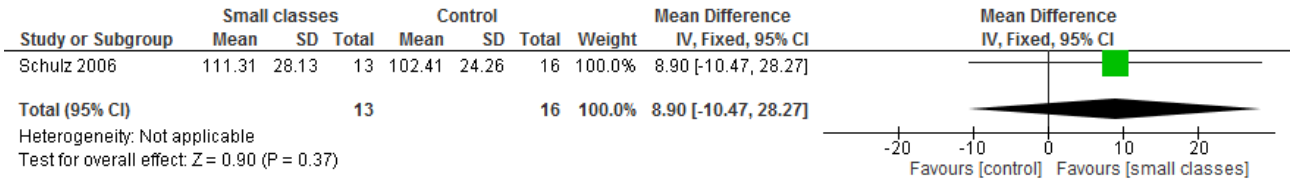
Comparison 5: Couple-focused classes versus standard care, outcome 1: Marital satisfaction women 6 months postnatal



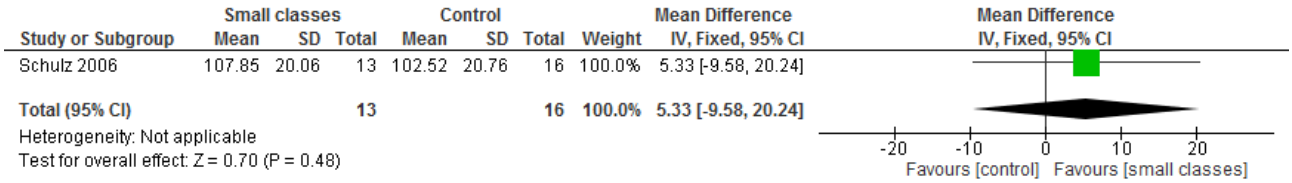
Comparison 5: Couple-focused classes versus standard care, outcome 2: Marital satisfaction men 6 months postnatal



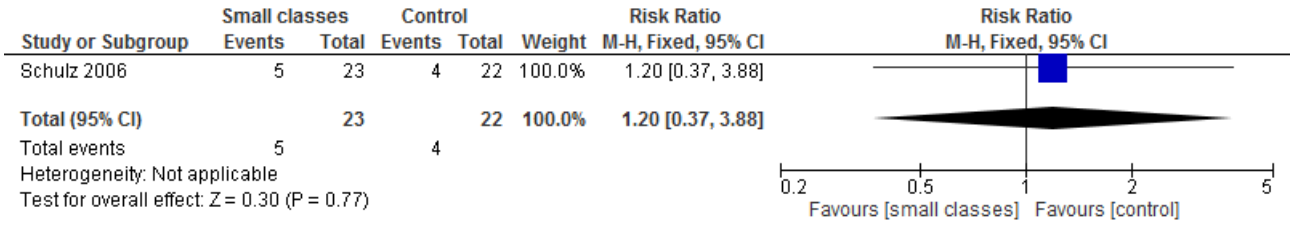
Comparison 5: Couple-focused classes versus standard care, outcome 3: Marital satisfaction women 5.5 years postnatal



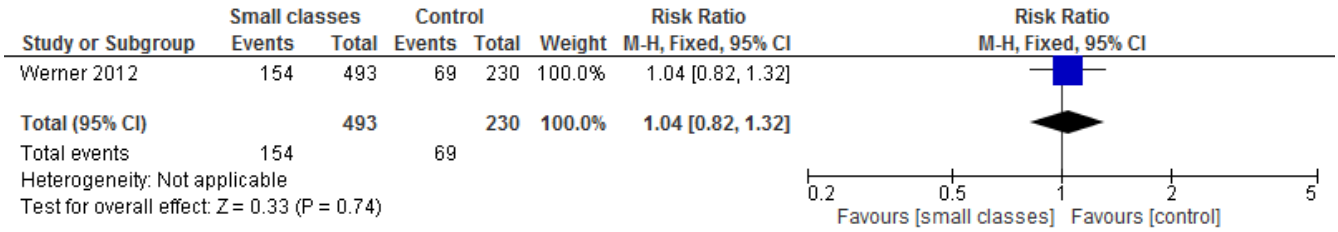
Comparison 5: Couple-focused classes versus standard care, outcome 4: Marital satisfaction men 5.5 years postnatal



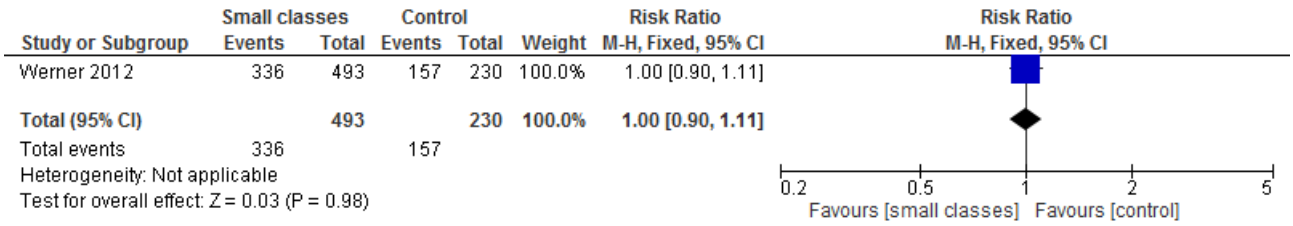
Comparison 5: Couple-focused classes versus standard care, outcome 5: Divorce/separation 5.5 years postnatal



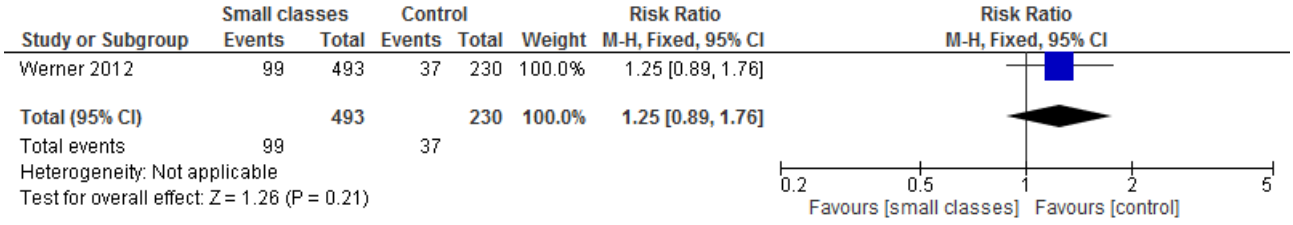
Comparison 6: Self-hypnosis versus standard care, outcome 1: Epidural analgesia



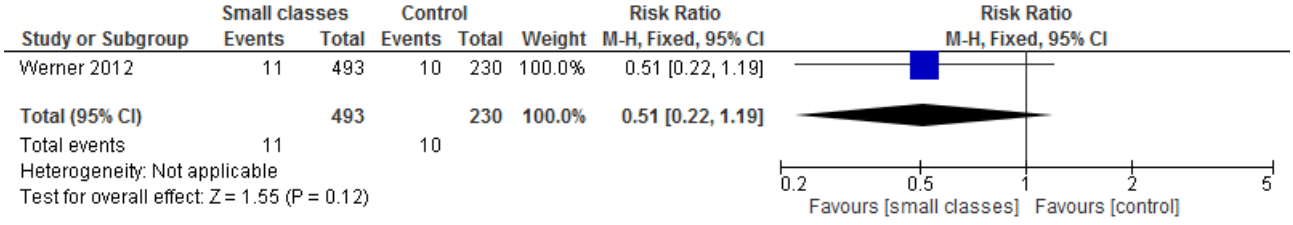
Comparison 6: Self-hypnosis versus standard care, outcome 2: Spontaneous delivery



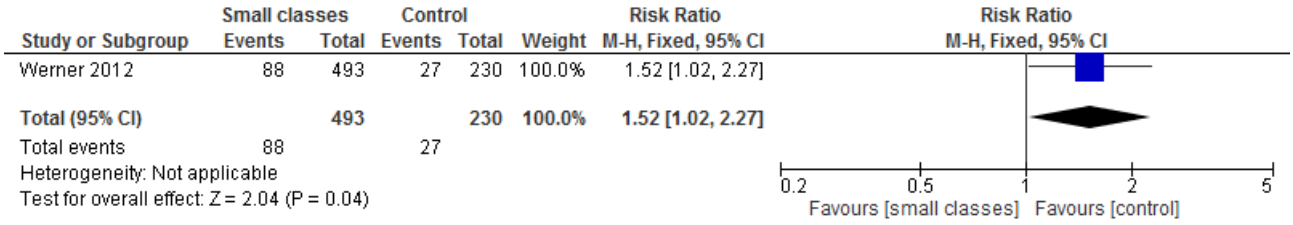
Comparison 6: Self-hypnosis versus standard care, outcome 3: Overall caesarean section



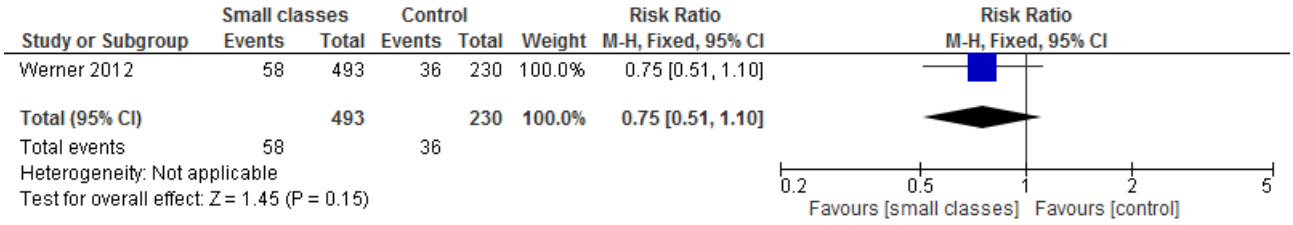
Comparison 6: Self-hypnosis versus standard care, outcome 4: Elective caesarean section



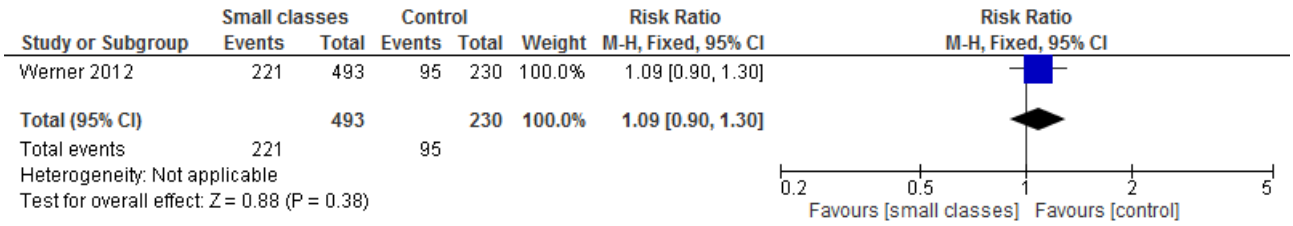
Comparison 6: Self-hypnosis versus standard care, outcome 5: Emergency caesarean section



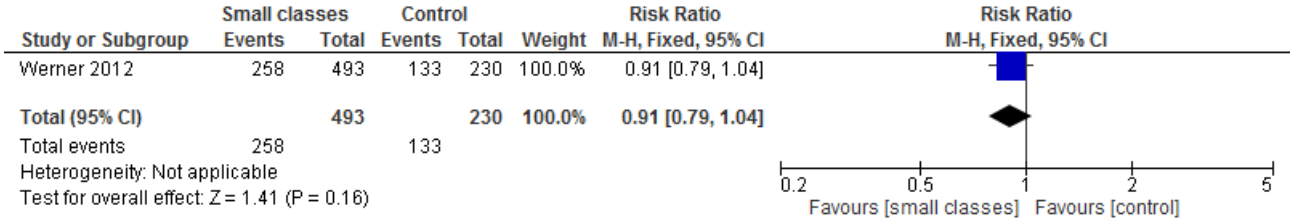
Comparison 6: Self-hypnosis versus standard care, outcome 6: Vacuum extraction



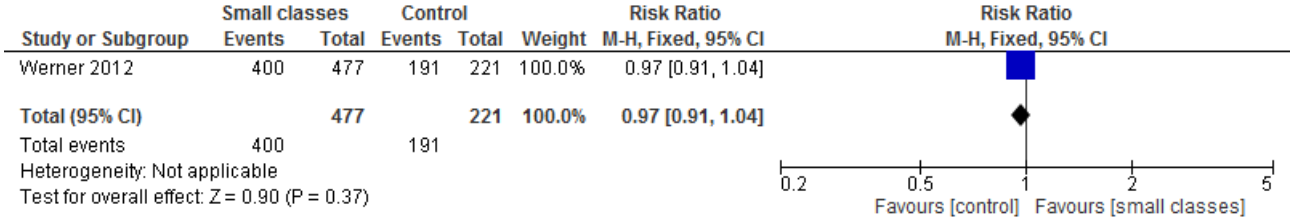
Comparison 6: Self-hypnosis versus standard care, outcome 7: Oxytocin augmentation



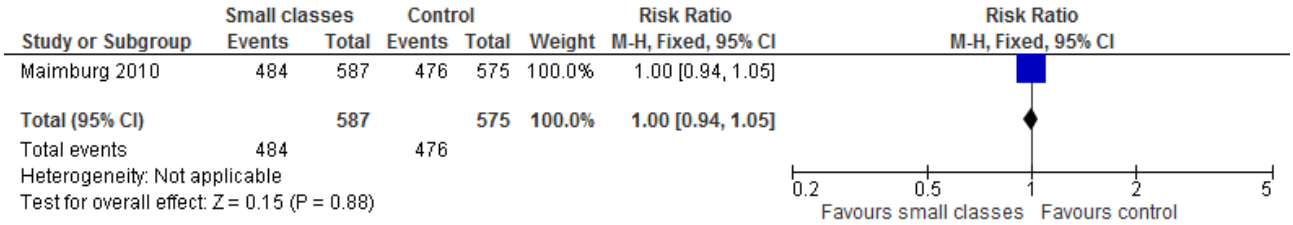
Comparison 6: Self-hypnosis versus standard care, outcome 8: Labor induction



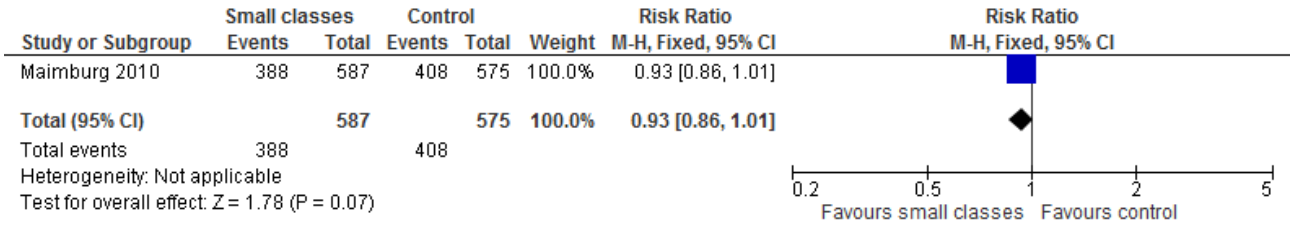
Comparison 6: Self-hypnosis versus standard care, outcome 9: Any breast feeding 4 months postnatal



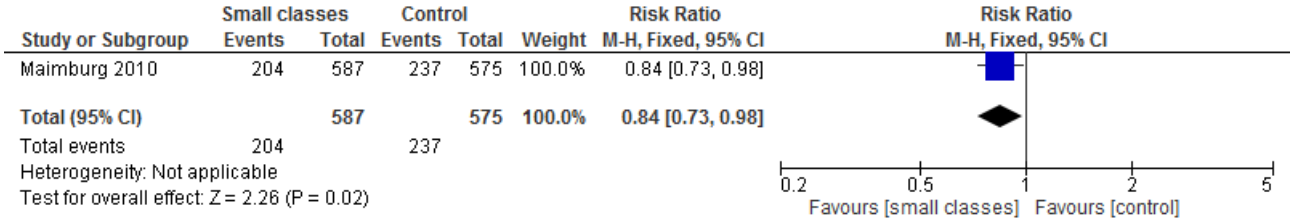
Comparison 7: General antenatal education classes versus standard care, outcome 1: Overall pain relief



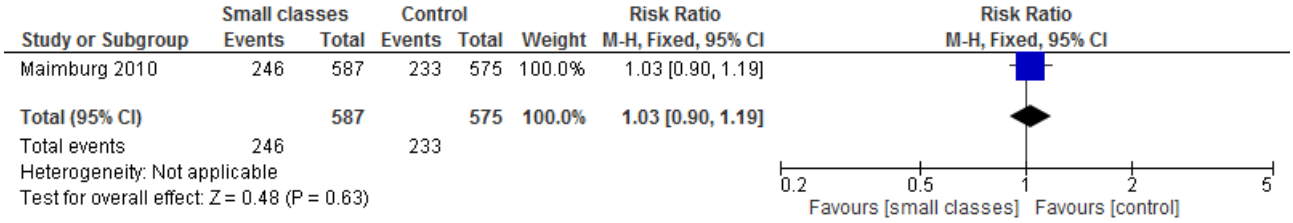
Comparison 7: General antenatal education classes versus standard care, outcome 2: Overall pharmacological pain relief



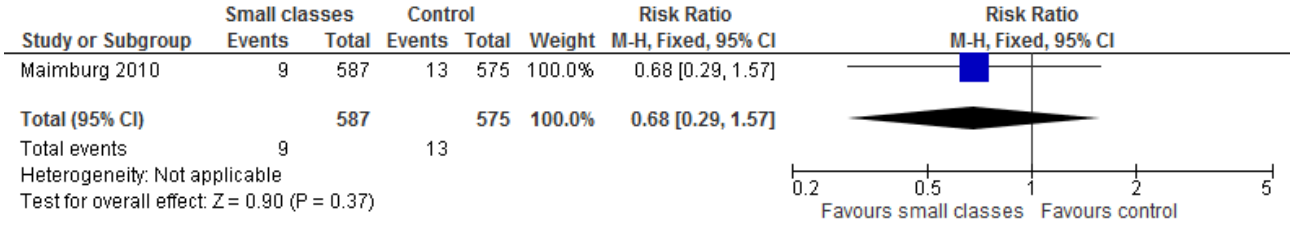
Comparison 7: General antenatal education classes versus standard care, outcome 3: Epidural analgesia



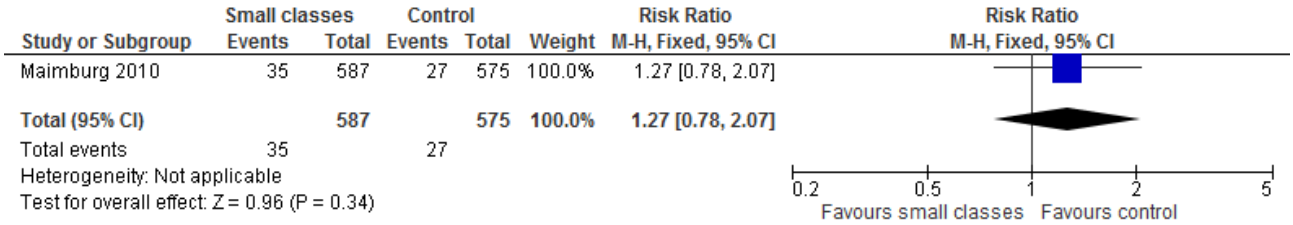
Comparison 7: General antenatal education classes versus standard care, outcome 4: Nitrous oxide/oxygen



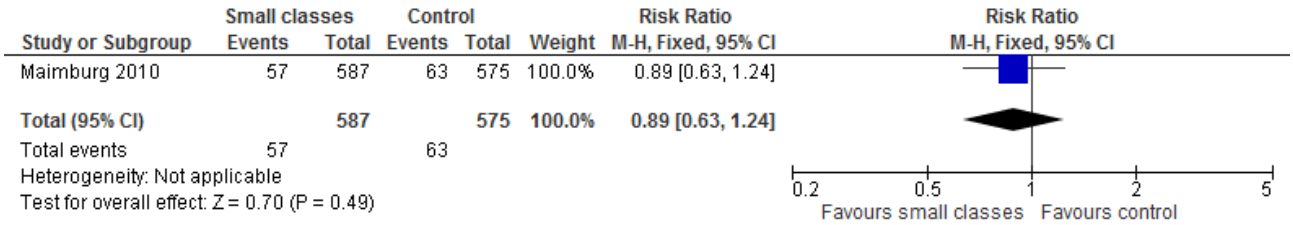
Comparison 7: General antenatal education classes versus standard care, outcome 5: Intramuscular morphine



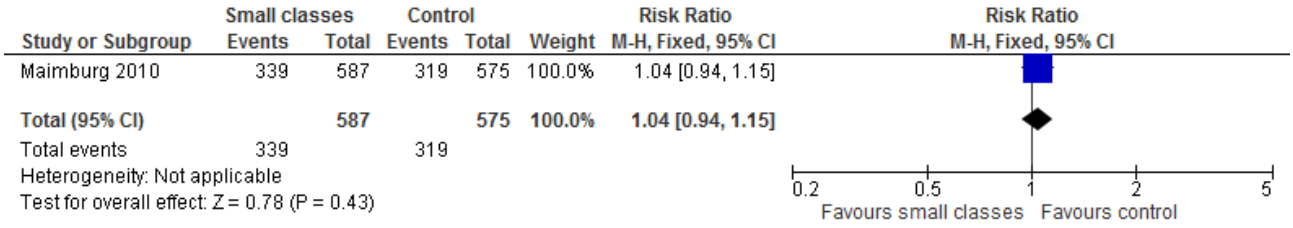
Comparison 7: General antenatal education classes versus standard care, outcome 6: Pudendal nerve block



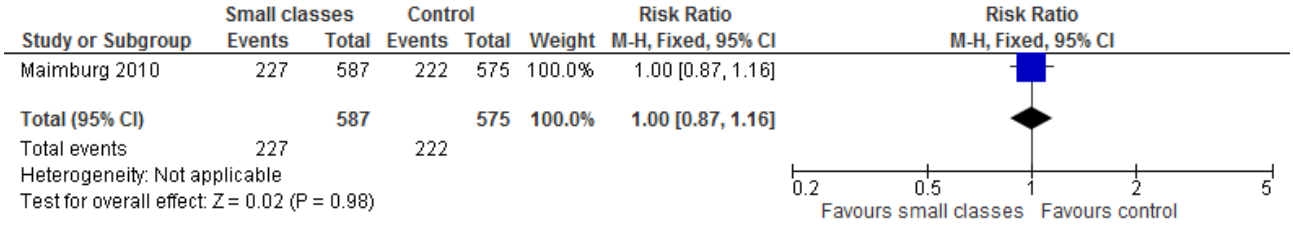
Comparison 7: General antenatal education classes versus standard care, outcome 7: Other pharmacological (primary halcion, codein, paracetamol)



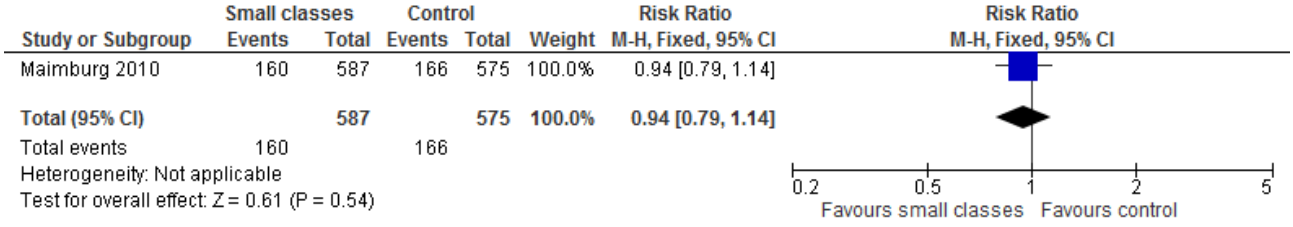
Comparison 7: General antenatal education classes versus standard care, outcome 8: Overall non-pharmacological pain relief



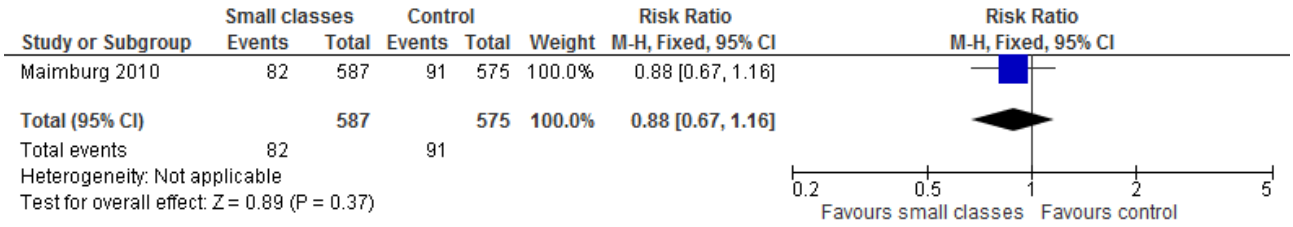
Comparison 7: General antenatal education classes versus standard care, outcome 9: Water immersion



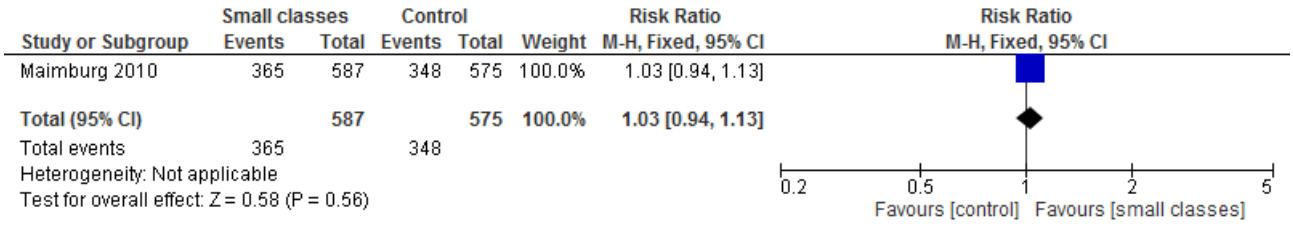
Comparison 7: General antenatal education classes versus standard care, outcome 10: Acupuncture



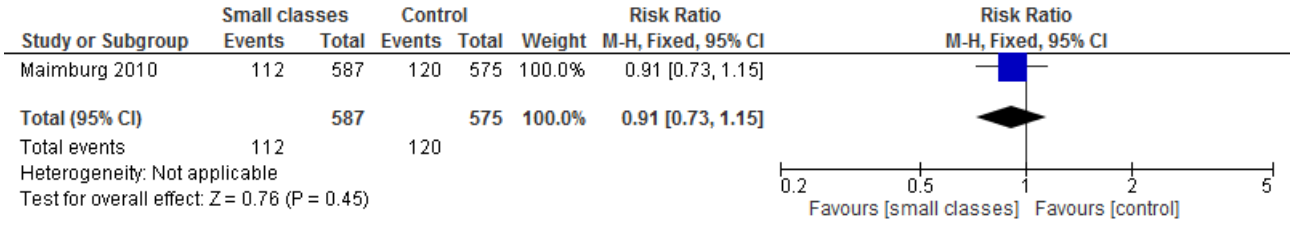
Comparison 7: General antenatal education classes versus standard care, outcome 11: Intracutaneous sterile water injection



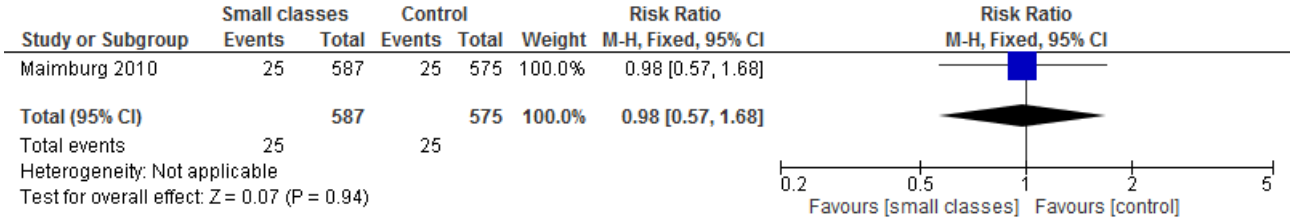
Comparison 7: General antenatal education classes versus standard care, outcome 12: Spontaneous delivery



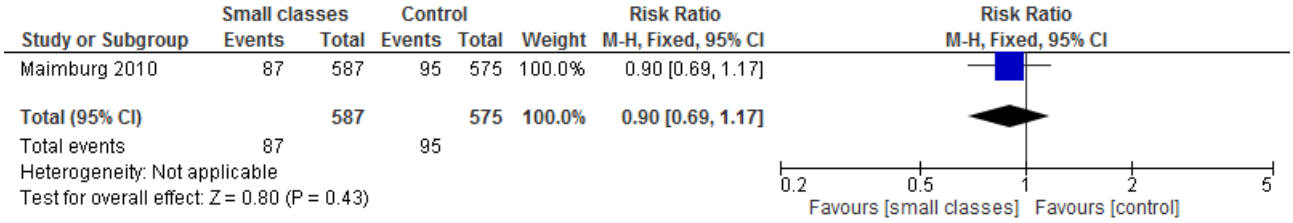
Comparison 7: General antenatal education classes versus standard care, outcome 13: Overall caesarean section



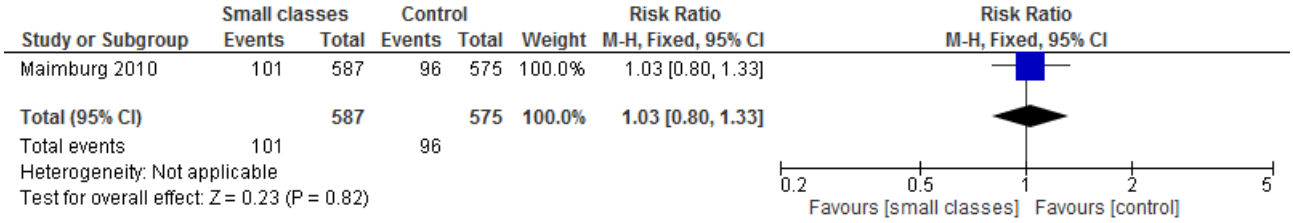
Comparison 7: General antenatal education classes versus standard care, outcome 14: Elective caesarean section



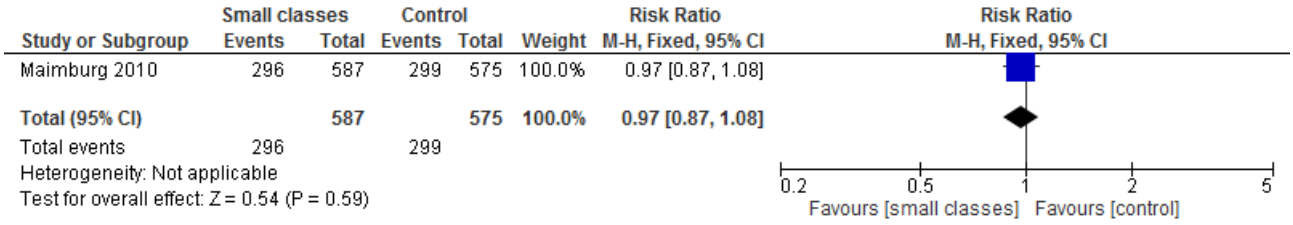
Comparison 7: General antenatal education classes versus standard care, outcome 15: Emergency caesarean section



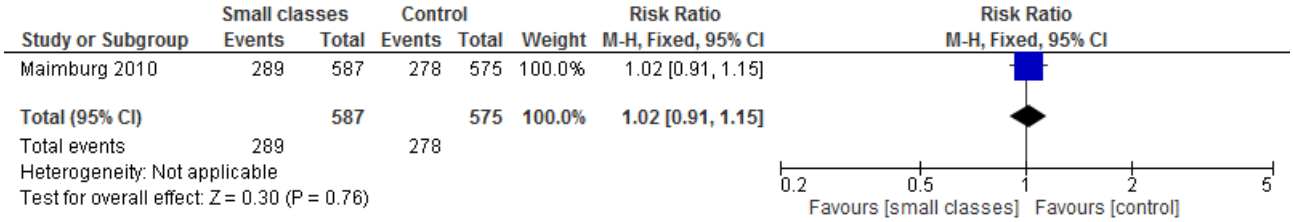
Comparison 7: General antenatal education classes versus standard care, outcome 16: Vacuum extraction



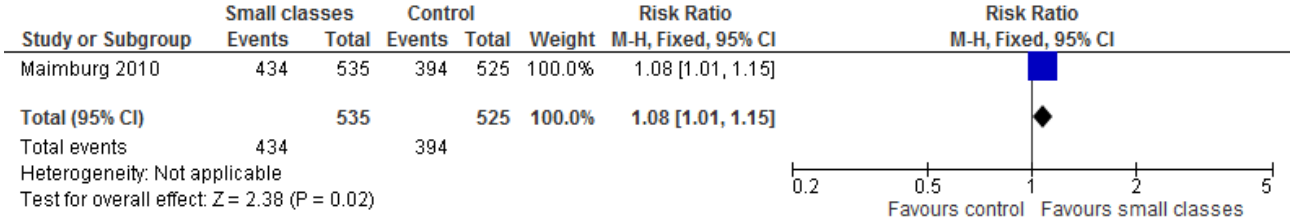
Comparison 7: General antenatal education classes versus standard care, outcome 17: Oxytocin augmentation



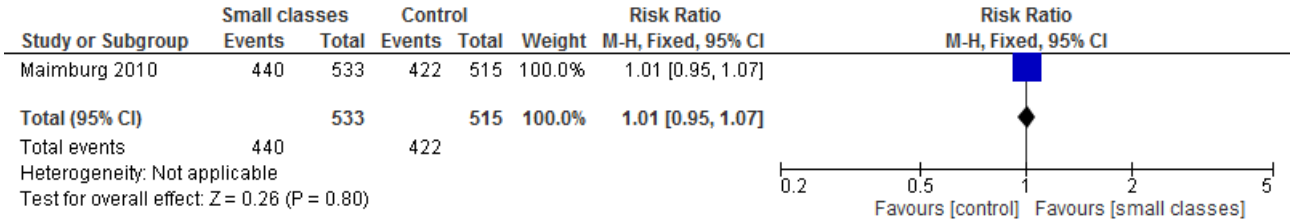
Comparison 7: General antenatal education classes versus standard care, outcome 18: Labor induction



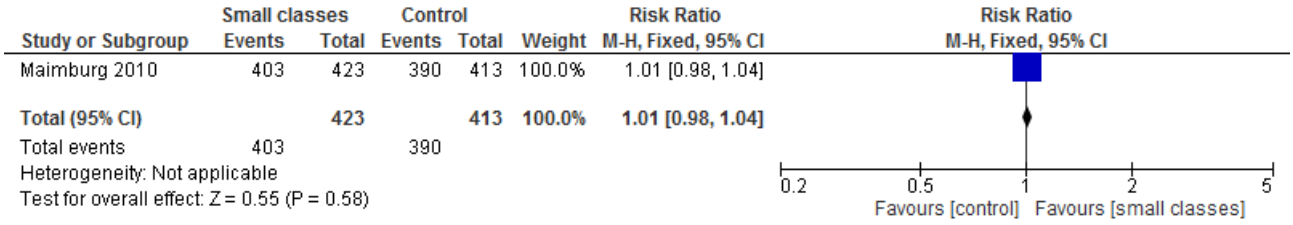
Comparison 7: General antenatal education classes versus standard care, outcome 19: Sufficient knowledge about breast feeding



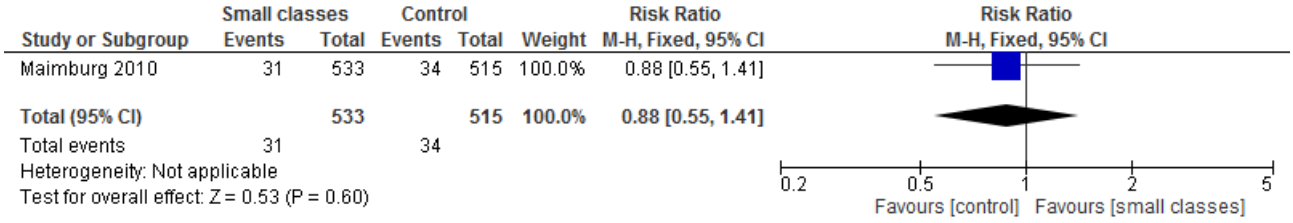
Comparison 7: General antenatal education classes versus standard care, outcome 20: Exclusive breast feeding 6 weeks postnatal



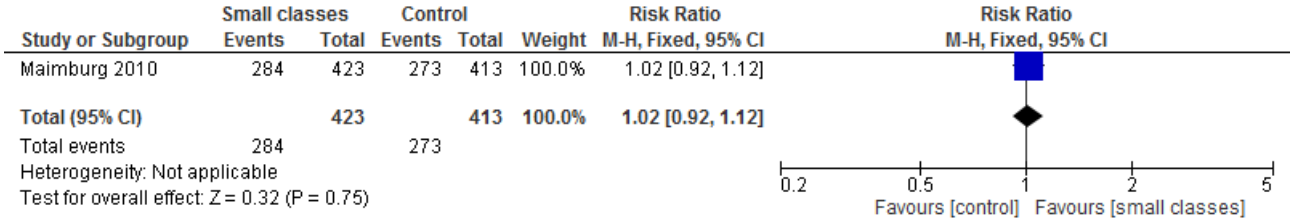
Comparison 7: General antenatal education classes versus standard care, outcome 21: Any breast feeding 6 weeks postnatal



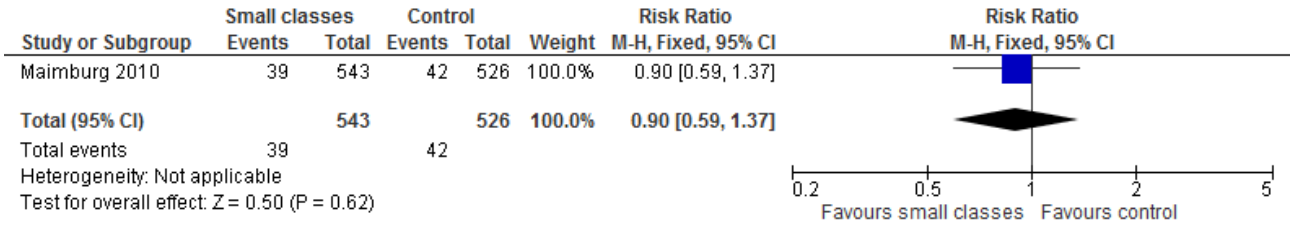
Comparison 7: General antenatal education classes versus standard care, outcome 22: Exclusive breast feeding 6 months postnatal



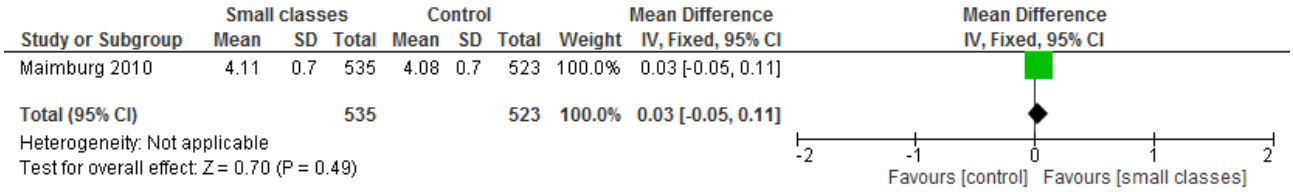
Comparison 7: General antenatal education classes versus standard care, outcome 23: Any breast feeding 6 months postnatal



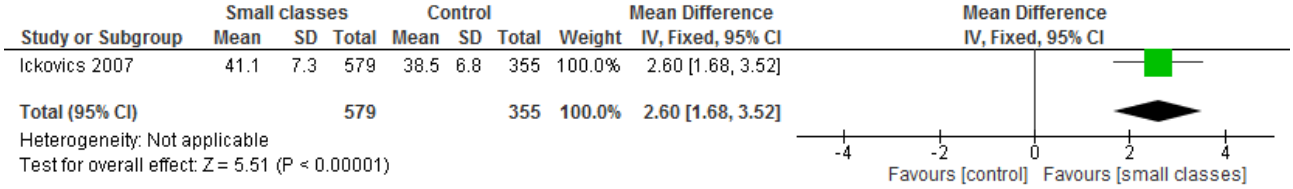
Comparison 7: General antenatal education classes versus standard care, outcome 24: Postnatal depression 6 weeks postnatal (EPDS)



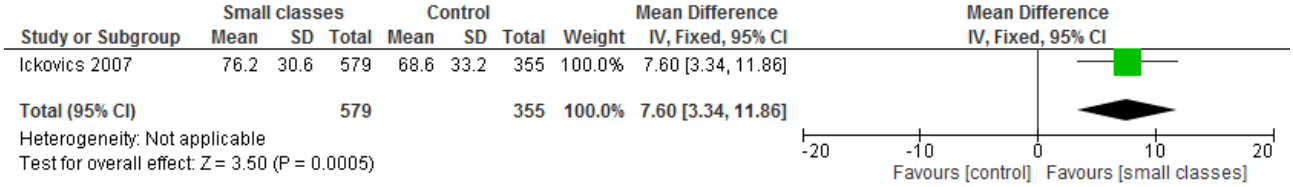
Comparison 7: General antenatal education classes versus standard care, outcome 25: Breast feeding self-efficacy 6 weeks postnatal



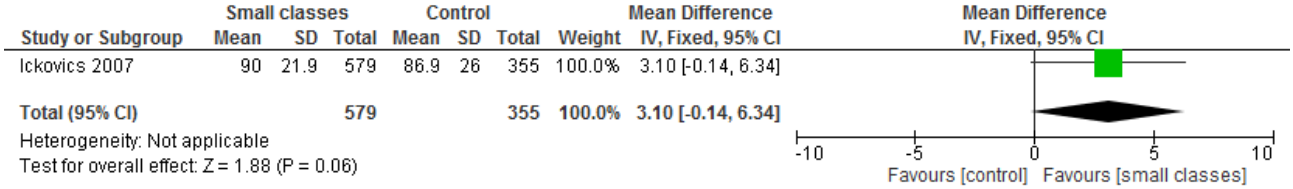
Comparison 8: Group prenatal care (20 hours) vs. individual prenatal care (2 hours), outcome 1: Prenatal and infant care knowledge



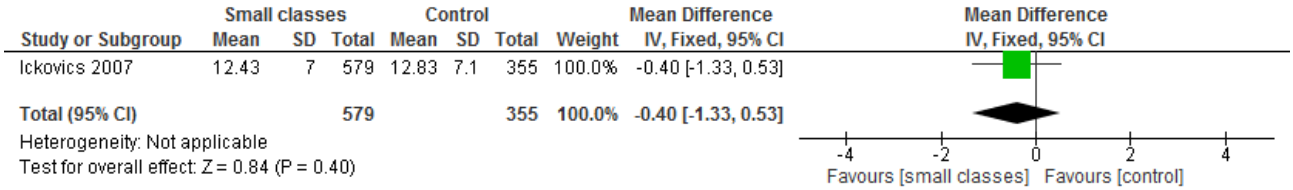
Comparison 8: Group prenatal care (20 hours) vs. individual prenatal care (2 hours), outcome 2: Readiness for labor and delivery



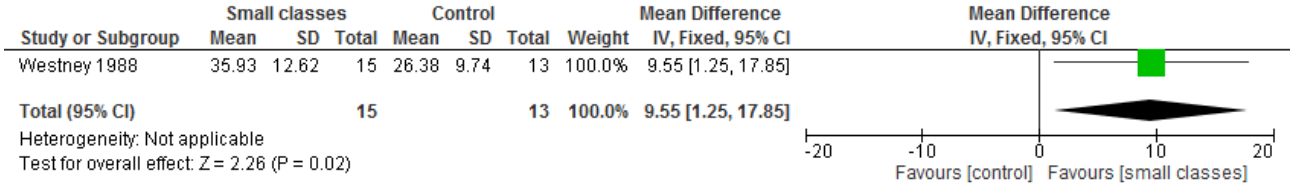
Comparison 8: Group prenatal care (20 hours) vs. individual prenatal care (2 hours), outcome 3: Readiness for infant care



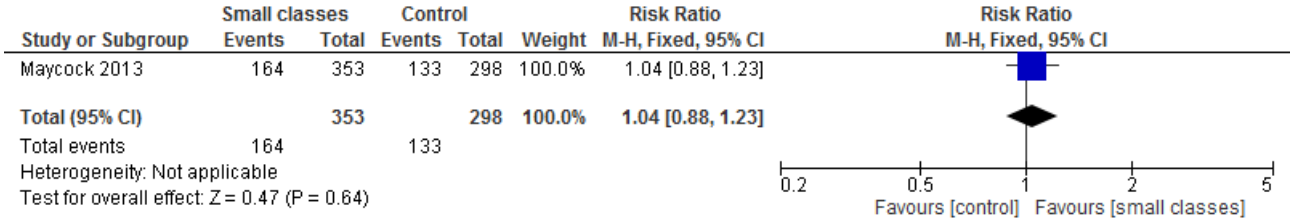
Comparison 8: Group prenatal care (20 hours) vs. individual prenatal care (2 hours), outcome 4: Prenatal distress



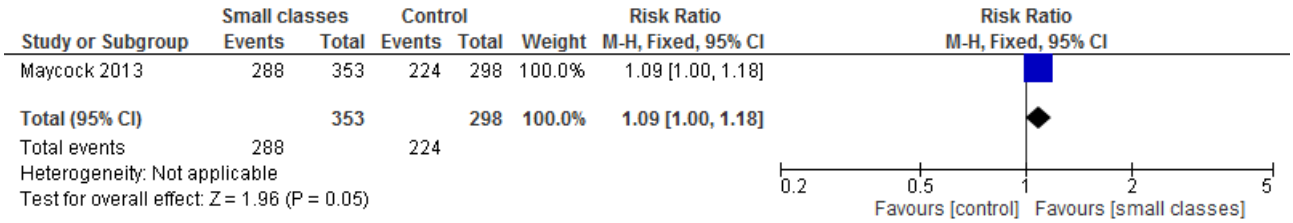
Comparison 9: Paternal education class versus standard care, outcome 1: Paternal knowledge



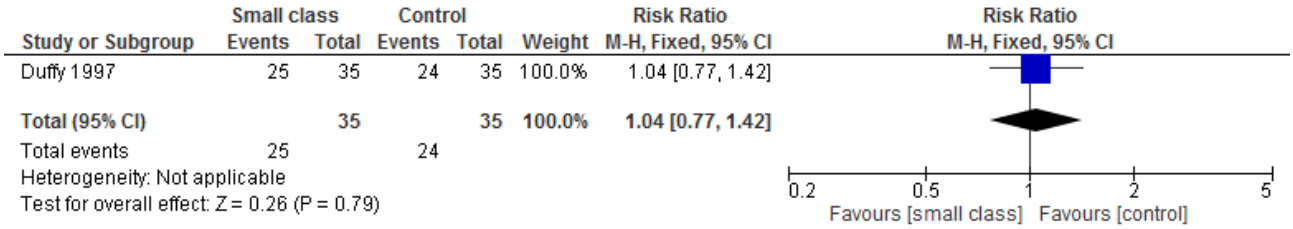
Comparison 9: Paternal education class versus standard care, outcome 2: Exclusive breast feeding 6 weeks postnatal



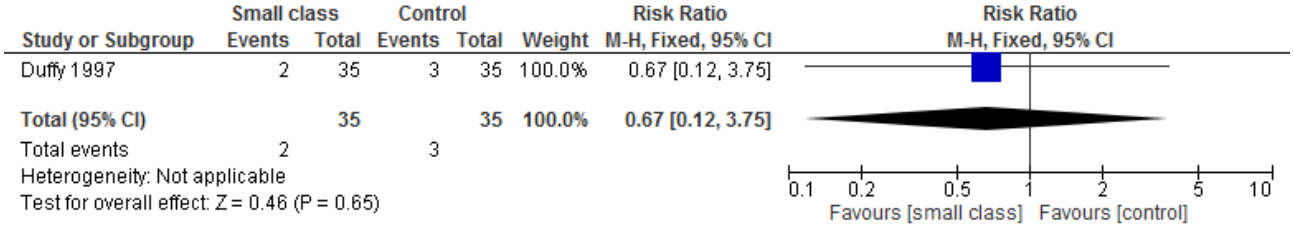
Comparison 9: Paternal education class versus standard care, outcome 3: Any breast feeding 6 weeks postnatal



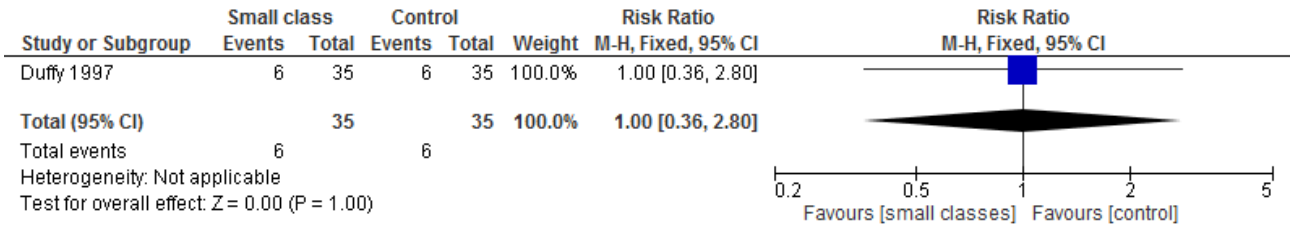
Comparison 10: Extra breast feeding sessions versus standard care, outcome 1: Spontaneous delivery



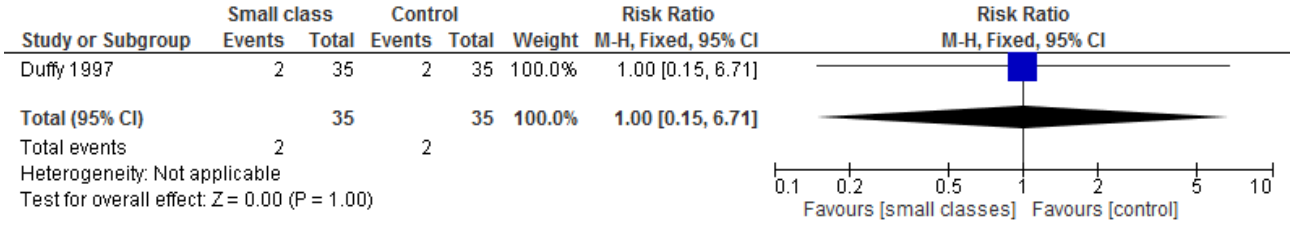
Comparison 10: Extra breast feeding sessions versus standard care, outcome 2: Overall caesarean section



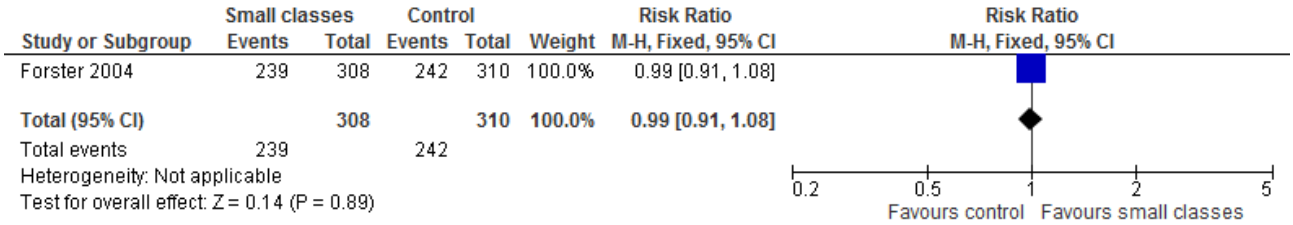
Comparison 10: Extra breast feeding sessions versus standard care, outcome 3: Vacuum extraction



Comparison 10: Extra breast feeding sessions versus standard care, outcome 4: Forceps

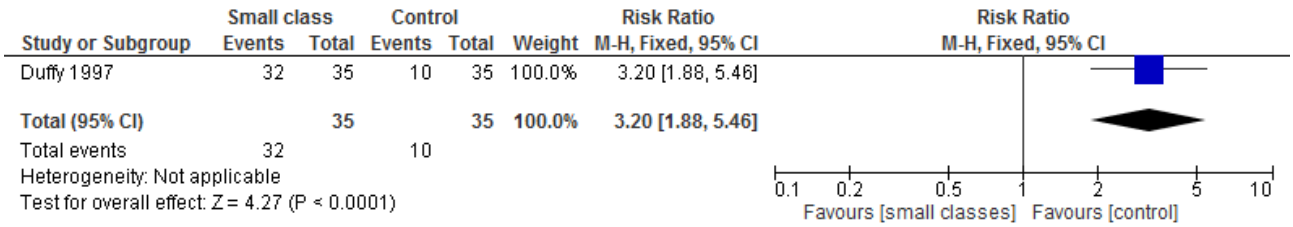


Comparison 10: Extra breast feeding sessions versus standard care, outcome 5: Breast feeding initiation – breast milk only

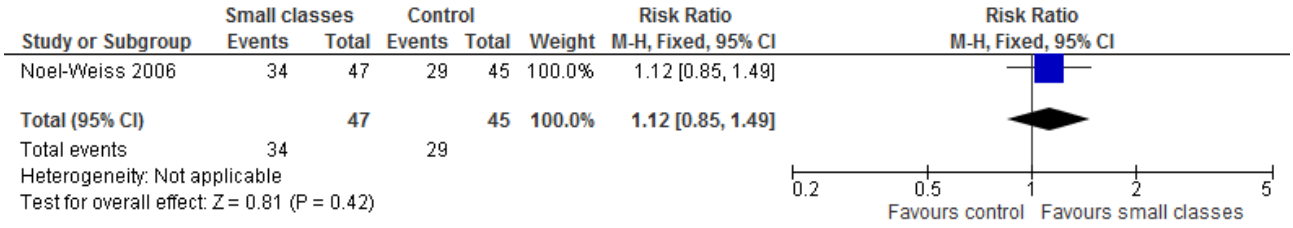


Comparison 10: Extra breast feeding sessions versus standard care, outcome 6: Breast feeding initiation – any breast milk

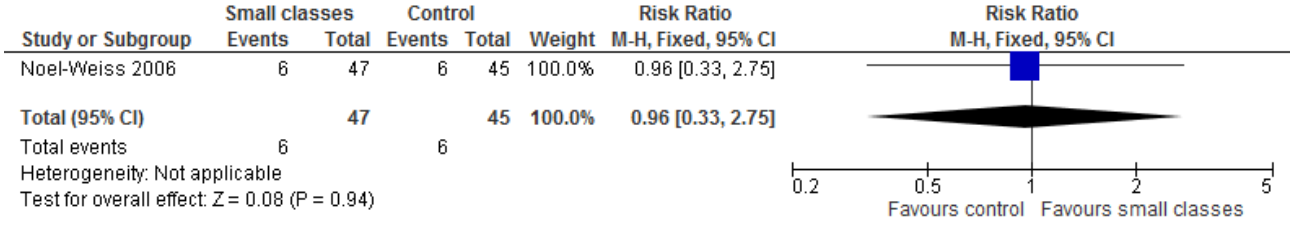
Comparison 10: Extra breast feeding sessions versus standard care, outcome 7: Exclusive breast feeding 6 weeks postnatal



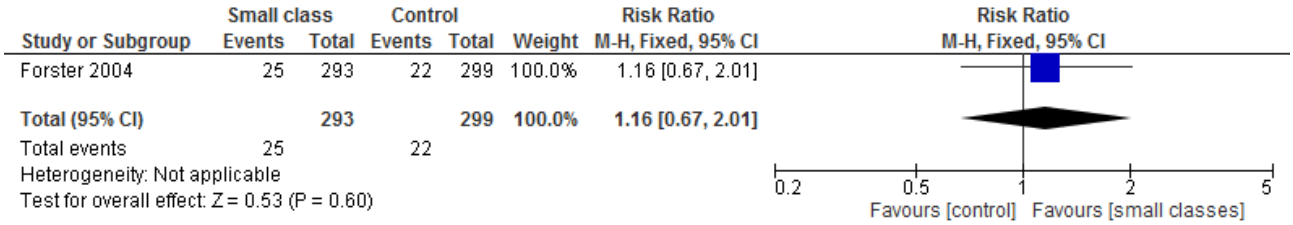
Comparison 10: Extra breast feeding sessions versus standard care, outcome 8: Breast milk only 8 weeks postnatal.



Comparison 10: Extra breast feeding sessions versus standard care, outcome 9: Any breast milk 8 weeks postnatal.



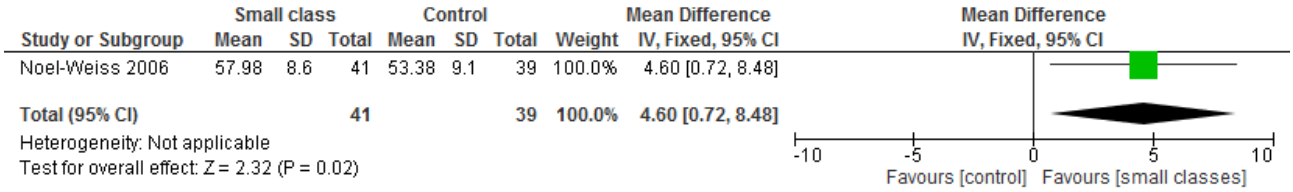
Comparison 10: Extra breast feeding sessions versus standard care, outcome 10: Exclusive breast feeding 6 months postnatal



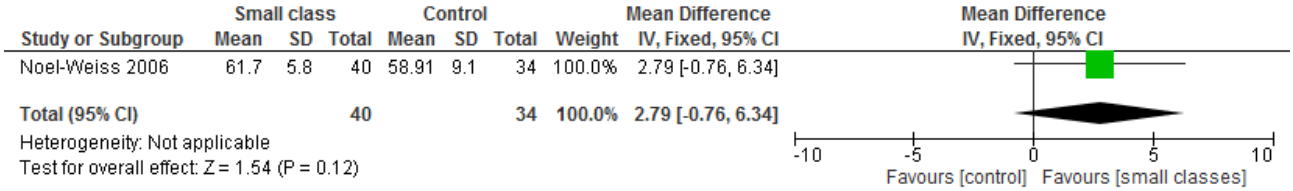
Comparison 10: Extra breast feeding sessions versus standard care, outcome 11: Breast milk only 6 months postnatal

Comparison 10: Extra breast feeding sessions versus standard care, outcome 12: Any breast milk 6 months postnatal

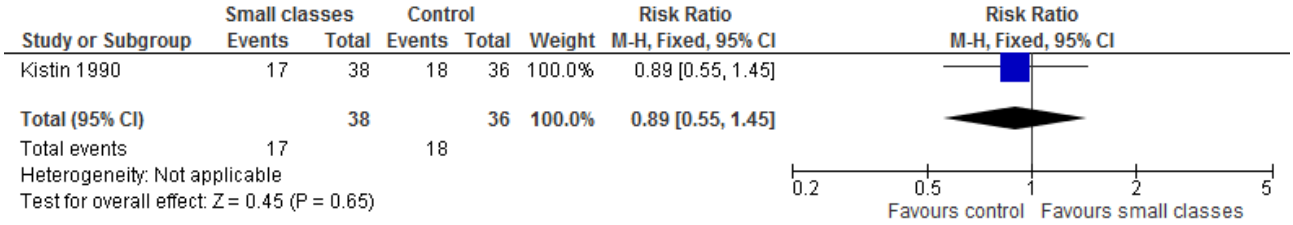
Comparison 10: Extra breast feeding sessions versus standard care, outcome 13: Breast feeding self-efficacy 4 weeks postnatal



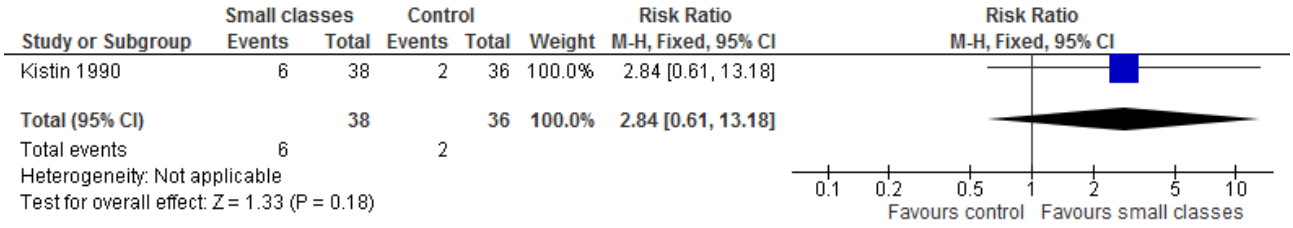
Comparison 10: Extra breast feeding sessions versus standard care, outcome 14: Breast feeding self-efficacy 8 weeks postnatal



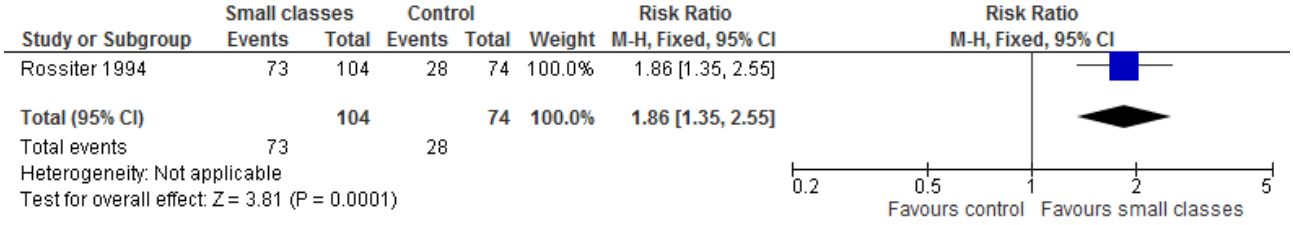
Comparison 11: Breast feeding classes versus one-to-one contact, outcome 1: Breast feeding initiation - one or more breast feedings per day



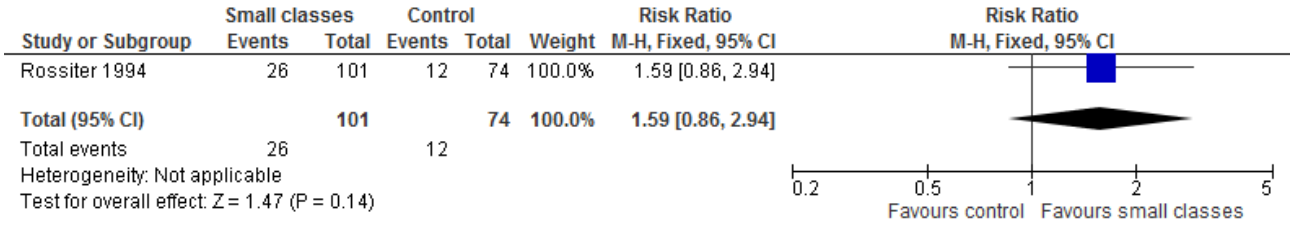
Comparison 11: Breast feeding classes versus one-to-one contact, outcome 2: Any breast feeding 12 weeks postnatal



Comparison 12: Breast feeding classes versus breast feeding and childbirth pamphlets, outcome 1: Breast feeding initiation - main source of nutrition



Comparison 12: Breast feeding classes versus breast feeding and childbirth pamphlets, outcome 2: Breast feeding as main source of nutrition 6 months postnatal



Characteristics of excluded trials

Main reason for exclusion

Not randomized trial	
Elliot 2000	Not a randomized controlled trial.
Turan 2003	Not a randomized controlled trial.
Outcome outside the scope of this review	
Mendelson 2008	Outcome: health behaviors, glycemic control, and neonatal outcomes among women with gestational diabetes.
Carter 1989	Outcome: congenital toxoplasmosis.
Haugland 2006	Outcome: pelvic girdle pain in pregnancy.
Morkved 2007	Outcome: lumbopelvic pain, sick leave and functional status.
Guelinckx 2010	Outcome: energy intake, physical activity, dietary habits and gestational weight gain in obese pregnant women.
Bogaerts 2013	Outcome: gestational weight among obese pregnant women.
Stafne 2012	Outcome: lumbopelvic pain.
Bonell 2013	Outcome: teenage pregnancy.
Miquelutti 2013	Outcome: lumbopelvic pain and urinary incontinence.
Hunter 2005	Outcome: prenatal diagnosis.
Hui 2012	Outcome: excessive gestational weight gain, physical activity, food intake.
Experimental condition outside the scope of this review	
Hunt 1976	Experimental condition: nutrition education.
Shapiro 2011	Experimental condition given as either a prenatal or postnatal workshop. More than half of the participants received the workshop postnatally.
Barakat 2008	Experimental condition: exercise.
Individual-based intervention	
Halonen 1985	Individual-based intervention aiming to reduce postpartum distress using relaxation training.
Leventhal 1989	Individual-based intervention aiming to increase positive childbirth experience.
Subramanian 2012	Individual-based intervention focusing on reducing behavioral and psychosocial risks.
Hayes 2001	Individual-based intervention aiming to decrease postnatal depression.
Mcleod 2004	Individual-based smoking and breast feeding intervention.
Kimber 2008	Individual-based music and massage intervention delivered during birth.
Spinelli 2013	Individual-based psychotherapy intervention focused on reducing postnatal depression.
Same dose of antenatal education in experimental and control condition	
Durham 1986	Only difference between conditions was music as a conditioning aid in childbirth.
Coffman 1994	Antenatal classes had different focus areas. Focus in experimental condition was partner support.
Wolfberg 2004	Antenatal classes had different focus areas. Focus in the experimental condition was breast feeding support.
Timpano 2011	Antenatal classes had different focus areas. The experimental condition was focused on OCD behavior – the control condition was focused more generally on anxiety.
Svensson 2009	Antenatal classes had different focus areas. More focus on parenting issues in the experimental condition with the aim of improving parental coping.
Bergstrom 2009	Antenatal classes had different focus areas. Focus in experimental condition was on natural birth and coping med labor pain.

Zimmermann-Tansella 1979	Antenatal classes had different teaching methods. The experimental education was focus on body sensations and the control condition focus was on lectures and discussions.
Kozinszky 2012	Antenatal classes had different focus areas. Focus in experimental condition was on psycho-education and psychotherapy for decreasing postpartum depression symptomatology.
Hawkins 2006	Only difference between conditions was a booklet and video segments on relationship deterioration.
Co-interventions in addition to antenatal education in small classes	
Sciacca 1995	Additional presents for breastfeeding, which was the main outcome.
Koniak-griffin 2000	Additional 17 home visits aiming at increasing health and social outcomes, and mother-child interaction.
Klerman 2001	Additional individual sessions several times during pregnancy. The intervention aimed at improving pregnancy outcomes and patients' knowledge of risks, satisfaction with care and behavior.
Doherty 2006	Additional home visits. The purpose of the intervention was to increase father involvement and skills with infants during the transition to parenthood.
Wambach 2011	Additional home visits and telephone counselling. The intervention focused on breastfeeding support and education.
Kieffer 2013	Additional home visits. The aim of the intervention was to reduce depressive symptoms among pregnant and early postpartum Latinas.
Halford 2010	Additional home visits. The intervention aimed to promote a positive transition to parenthood.
Turan et al. 2001	Additional individual telephone consultations. The paper is a summary of three studies on methods for including men in antenatal education.
Inadequate information for analysis	
Olenick 2010	No description of control condition. The aim of the intervention was to improve breastfeeding outcomes.
Richter 2012	No description of control condition. The aim of the intervention was to reduce stress in pregnant women in high risk of stress.
Wolfson 1992	No information on the number of participants providing outcome data at each time point by group. This trial studied the effect of parent training on infant sleeping patters, parents' stress, and perceived parental competence.
Lavender 2005	No information on number of participants providing outcome data at each time point by group. This trial evaluated the effect of a breastfeeding intervention on breastfeeding duration.
Schachman 2004	No standard deviations on means. The aim of the intervention was to increase prenatal and postpartum maternal role adaptation.
Midmer 1995	No standard deviations on means. The aim of the intervention was among other things to increase marital adjustment, and postpartum adjustment.
Matthey 2004	Results are presented as stratified analyses – no raw effect is presented. The aim of the intervention was to increase postpartum psychosocial adjustment.
Kermeen 1995	Only F statistics and p-values for the comparison between groups are presented - no effect size measures. The aim of the intervention was to lower the potential negative effect of becoming a parent.

Paper III

Sexual & Reproductive Healthcare 4 (2013) 121–126

Contents lists available at ScienceDirect

Sexual & Reproductive Healthcare

journal homepage: www.srhcjournal.org

Group-based antenatal birth and parent preparation for improving birth outcomes and parenting resources: Study protocol for a randomised trial

Vibeke Koushede [a,*], Carina Sjöberg Brixval [a], Solveig Forberg Axelsen [a], Jane Lindschou [b], Per Winkel [b], Rikke Damkjær Maimburg [c,d], Pernille Due [a], and the NEWBORN trial group

[a] National Institute of Public Health, University of Southern Denmark, Denmark
[b] The Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Denmark
[c] School of Public Health, Section of Epidemiology, Aarhus University, Denmark
[d] Aarhus University Hospital, Department of Obstetrics and Gynaecology, Skejby, Denmark

ARTICLE INFO

Article history:
Received 14 June 2013
Revised 14 August 2013
Accepted 16 August 2013

Keywords:
Randomised trial
Intervention research
Antenatal preparation
Pregnancy
Birth
Parenthood

ABSTRACT

Objectives: To examine the efficacy and cost-effectiveness of group based antenatal education for improving childbirth and parenting resources compared to auditorium based education.
Methods: Participants: 2350 Danish pregnant women and their partners ⩾18 years old, recruited before 20 + 0 gestational weeks.
Population-based individually randomised superiority trial with two parallel arms: Four sessions of birth and parent preparation in small groups (experimental group); two lectures in an auditorium (control group).
Data is collected by (1) questionnaires at baseline (≈18 weeks of gestation), 37 weeks of gestation, 9 weeks-, 6 months-, and 1 year post-partum, (2) the hospital obstetric database, (3) national registers. Primary outcome: use of epidural analgesia. Secondary outcomes: stress, parenting alliance; explorative outcomes: depressive symptoms, use of health care services, self-efficacy, well-being, family break-ups. Analyses will be intention-to-treat as well as per protocol. Process evaluation will be conducted using questionnaires and qualitative interviews. The incremental societal cost of the intervention will be computed and compared to the measured outcomes in a cost-effectiveness analysis.
Conclusion: To the best of our knowledge this is the largest well-designed randomised trial of its kind to date. The trial will bring much-needed evidence for decision makers of the content and form of antenatal education.

Introduction

The majority of prospective parents look to antenatal education to gain information on issues such as decision making about and during labour, infant and postnatal care, breastfeeding, and parenting skills [1]. However, the form and content of antenatal education has been sensitive to opinions and trends and has undergone many changes without specific evidence on its effects on relevant outcomes for parents and children.

Today, the main focus of many antenatal classes is birth and breastfeeding; while information on parent–child attachment and psychosocial aspects that relate to couple- and parenthood are generally not covered [2–5], although studies suggest that parents need this information [6]. Further, many antenatal classes are conducted in large auditoriums. It has been argued that information transfer should no longer be the focus of antenatal education. Experts suggest that educators need to become facilitators and emphasis should be shifted from the educator to the learner. Furthermore, that individuals need to interact with fellow learners and the learning environment in order to learn and obtain new competencies [7].

Previous studies of antenatal education have been difficult to interpret, mainly because of limitations in study design, high risk of bias, and small sample sizes [1]. Further, only few randomised trials have been conducted [1]. A Cochrane review of all randomised trials about individual or group antenatal education for childbirth or parenthood from 2007 concludes that the effects remain largely unknown [1]. Since then only few well-conducted randomised trials have been carried out. These suggest a positive effect of attending antenatal education, e.g. on the birth process

* Corresponding author. Address: National Institute of Public Health, University of Southern Denmark, Øster Farimagsgade 5A, 2nd Floor, 1353 Copenhagen, Denmark. Tel.: +45 6550 7864.
E-mail addresses: vibe@niph.dk, vibe@si-folkesundhed.dk (V. Koushede), cabr@niph.dk (C.S. Brixval), sfax@niph.dk (S.F. Axelsen), JL@ctu.dk (J. Lindschou), pwinkel@ctu.dk (P. Winkel), rmai@soci.au.dk (R.D. Maimburg), pdu@niph.dk (P. Due).

1877-5756/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.srhc.2013.08.003

[8] and on parenting self-efficacy [9]. However, the effect of antenatal education on the overall acquisition of knowledge, on parent's ability to care for infants and to making psychological and social adjustments in the transition to parenthood is still largely unknown.

Cochrane reviews of randomised trials on parent training programmes suggest that parenting programmes have a potential role to play in the promotion of mental health [10,11]; however, randomised trials on antenatal parenting programmes are scarce. Most current evidence from randomised trials addresses the use of parenting programmes, as part of secondary, high-risk approaches to prevention. However, it has been argued that they would be more effective if delivered as part of a population-based approach [11,12]; in this way they are offered to all parents to prevent problems and promote child and parent health.

To date, it is therefore unknown if: (1) antenatal preparation in small groups is superior to auditorium-based education; (2) which elements the classes should encompass to meet the parents' needs today; and (3) what the cost-effectiveness of antenatal preparation in small groups is compared to large-scale education in auditoriums.

Over the past years, Danish antenatal education has gradually moved away from large-scale auditorium-based education to antenatal birth and parent preparation classes in small groups for all expectant parents. In the Capital Region of Denmark antenatal classes are, however, still offered as lectures in hospital auditoriums with a minimum of interaction with the audience, although this region is planning to implement birth and parent preparation in small groups. The Capital Region of Denmark is therefore an ideal setting for a randomised trial of group-based versus auditorium-based antenatal education.

All healthcare systems have limited resources, and it is therefore important to develop a research-based up-to-date antenatal preparation programme and investigate its efficiency and cost-effectiveness prior to implementation.

Objective

The primary objective of the NEWBORN trial is to compare birth outcomes, parenting resources, health and thriving, and use of healthcare services in families enrolled in a research-based standardised antenatal birth and parenting programme in small groups with those allocated to auditorium-based education.

Process evaluation: As the degree of implementation of a programme is crucial to its ability to achieve any effect, we aim to ensure careful implementation of the programme. We will conduct a thorough process evaluation highlighting enabling factors and barriers to the implementation.

Cost-effectiveness analysis: Finally, the incremental societal cost of the intervention will be calculated and compared to the measured outcomes in a cost-effectiveness analysis.

Methods

Trial design

Population-based individually randomised superiority trial with two parallel arms: Four sessions of birth and parent preparation in small groups of 6–8 couples (experimental group); compared to two lectures in an auditorium on birth and breastfeeding, with participation of up to 250 people (control group).

Setting

The trial is conducted at the Department of Obstetrics and Gynaecology, Hvidovre Hospital (HH) in the Capital Region of Denmark. Denmark ensures tax-financed, free, equal access to medical health care.

Participants

Women will be enrolled from $10 - 20 + 0$ weeks of gestation. Inclusion criteria are expectant women, $\geqslant 18$ years old at enrolment, due to give birth at Hvidovre Hospital, able to speak and understand Danish, and being legally able and actually providing signed consent. The women's partners are also invited to participate.

The women will receive a written invitation to participate in the trial prior to their first visit to the birth clinic. Baseline data will be collected from the women and their partners when they accept the invitation and before randomisation. Oral and written information will be provided and the women and their partners will be randomised to the experimental group or the control group if she signs and returns the informed consent form.

Randomisation

Central randomisation will be performed using the Copenhagen Trial Unit's Online Randomisation system. A project employee will perform individual web-based randomisation according to a computer-generated allocation sequence with a varying block size concealed to the investigators. All citizens in Denmark have a unique personal identification number; the randomisation programme is set up to confirm the existence of the unique personal identification number. The allocation of participants will be 1:1 to the experimental group and the control group, respectively, stratified by vulnerability as defined by the birth site (vulnerable women: women with a previous or actual psychiatric illness, with an actual life crisis, who are victims of violence or are socially strained, versus women who are not vulnerable), and parity (nulliparous women versus multiparous women).

Intervention

The trial will have two intervention groups:

(1) The "NEWBORN" programme (experimental group). A research-based and theoretically founded birth and parenting programme developed in collaboration with midwives, health care visitors, psychologists and family therapists, parents, and leading national and international researchers and clinicians in this field (described in more detail below).
(2) Standard care (control group). The pregnant woman and her partner are offered two antenatal lectures on birth and breastfeeding in an auditorium with participation of up to 250 people.

The NEWBORN programme includes short verbal presentations from the group facilitator, individual exercises, short film presentations, time for discussions and reflection. Parents are given homework in the form of minor exercises in preparation to each session. Educational subjects are: the transition to parenthood; couple communication; birth; breastfeeding; and taking care of a newborn. A patient-network website has been created as a supplement to the sessions. The programme is focused on parenting resources important to the birth process, parenting and mental health, and that appear amenable to change, i.e.: social support, parenting alliance, cognitive coping, and parenting skills:

- *Social support:* formal and informal, emotional, informational and instrumental. Groups of 6–8 couples are offered three times 2.5 h sessions during pregnancy and one session 5 weeks postpartum. The groups are composited to enable participants establish relations with other expectant parents in their local area. Sessions are led by a midwife and the postnatal session will include a health visitor. A patient-network website enables parents to gain further information, communicate with other parents and consult online with a midwife and a health visitor.
- *Parenting alliance:* adding a component supporting the couples in the transition to parenthood and couple communication.
- *Cognitive coping:* embedding sources of self-efficacy into programme content and delivery, and by creating an environment which enables parents to discuss feelings and concerns, enhances their awareness of own resources, problem-solving strategies, and future challenges in parenting and emotional regulation
- *Parenting skills:* increasing information and exercises with feedback, e.g. on recognising signs and symptoms of thriving in the newborn, couple communication, etc.

The approach aims at strengthening relationships and improving information and problem solving skills for expectant parents in order to ease birth and the transition to parenthood.

To maximise the potential for population uptake classes have been established at three local midwifery sites. A comprehensive guide and education material for course facilitators has been developed, and facilitators, i.e. midwives and health visitors are trained at 1-day workshops. The framework for the classes is based on an estimate of adequate time allocated to each subject, and what service providers deem a sustainable service.

Session 1 (approximately 25 weeks gestation):

- The transition to parenthood – new roles and responsibilities.
- Common changes in the relationship during and after pregnancy.
- Couple communication.
- Meaning of own childhood when becoming parents.

During this session participants are introduced to one another and to the outline of the entire programme. The midwife invites the parents to think about and since discuss their expectations of the greatest joys and greatest challenges of parenthood. The midwife informs the parents about common changes and challenges in the relationship during pregnancy and after birth, and the importance of good communication. A short film is shown teaching the parents about good communication skills. The film is developed by the Danish Centre of Family Development and is inspired by PREP [13]. The film is followed by practical couple-communication exercises. Women attending the classes alone either pair up or conduct the exercise with the midwife. In short the exercise entails one person listening actively and without interrupting while the other person describes thoughts and feelings about a certain topic. Afterwards the person listening is encouraged to acknowledge what she/he has heard before changing roles. This communication exercise is used throughout the entire programme covering different topics e.g. expectations of parenthood, labour, the relationship after birth, etc. The aim is to try to understand the other person's perspective before trying to be understood. The importance of one's own childhood when becoming a parent is also a topic in this session [14]. Participants are asked to think back to their own childhood, how they were raised, their parents' parenting style (e.g. warm and affectionate, strict), traditions, etc. Afterwards they are to consider things they would like to carry forward into their own parenting as well as things they might want to do differently. Finally the couple discuss the topic using the communication technique they have been taught. The aim is to start a thought process. As there is not sufficient time for long in depth discussions participants are encouraged to carry on the discussions at home. During the first break participants who wish to do so are asked to write down their contact details so that they can be shared. Throughout all the sessions the midwife has a facilitating role helping discussions along if needed and commenting where appropriate. At the end of all sessions participants are asked to consider and write down the most important take home points of the session.

Suggested homework: seven short informational film clips (duration between 2 and 7 min) on the first signs of labour, the time at home in early labour, birth, when there is a need for obstetric intervention, and pain relief.

The films, exercises, and written information on session topics are available on the network website.

Session 2 (approximately 33 weeks gestation):

- Expectations in relation to birth.
- The normal course of labour.
- When there is a need to intervene in labour.
- Pain relief and coping strategies.
- Partner support during labour.

The aim of this session is providing participants with information, and enhancing their existing knowledge and understanding of the normal course of labour, pain-relief, and what might be expected if there is a need for obstetric intervention. After a short presentation by the midwife the couples discuss their hopes and expectations for labour and birth using the communication framework – they are asked to consider their individual resources and action competencies in relation to increasing the likelihood of obtaining their wishes. Also the couples are asked to discuss how they might support one another during labour and birth – using practical examples. For the topic on pain relief the women are asked to discuss their thoughts and previous experiences with coping with pain and physical and mental strain – what did they do, what helped them, can they use any of these strategies during labour? Next they discuss their thoughts and knowledge on various methods of pain relief. Meanwhile the men discuss their thoughts and feelings about their role during labour and birth. Plenary discussions and summing up thoughts and ideas are used so that participants can learn from and be inspired by one another. Vicarious learning and feedback are considered important in relation to self-efficacy [15].

Suggested homework: participants are encouraged to ask women in their social network about their breastfeeding experiences, and read a pamphlet that is handed out on breastfeeding [16].

Session 3 (approximately 35 weeks gestation):

- Feeding a newborn.
- Interpreting the newborn's signs, symptoms and behaviour.
- Taking care of a newborn.
- Mood swings and postnatal depressive symptomatology.

Participants discuss wishes for feeding their newborn and feeding experiences in their networks in small groups. The midwife then talks about how expectations, support, and the advice received from family and friends may affect e.g. breastfeeding intention and perseverance in the case of difficulties. Bearing the breastfeeding experiences of individuals in their social networks in mind (preparation for this session), participants are encouraged to consider who it might be most helpful to seek breastfeeding support and advice from if necessary.

Cards with a variety of breastfeeding topics are spread out on the table, and participants are asked to pick a topic that they wish

to hear more about, and tell the group why they have picked the chosen topic. Topics include e.g. how to tell that the baby is getting enough milk, positioning, importance of partner support, feeding patterns, breast engorgement, etc. There are certain topics that the midwife is told to cover regardless of whether it has been picked by a participant or not e.g. how to tell the baby is getting enough milk. The pilot study showed that participants are likely to choose a topic they already know something about in order to receive verification and feedback from the midwife and from the other participants (unpublished data) – this may help increase self-efficacy.

The midwife gives information and shows short film clips on baby cues and sleep patterns. The importance of communicating with the newborn is underlined. Information on the prevention of cot death is given. Next the initial time at home with a newborn and the importance of social networks for emotional and practical support is discussed (the group is considered a potential supportive social network). Participants are given an exercise where they are asked to fill in a list of expected daily activities after the baby is born. Afterwards they compare their list with their partner's or that of another group member before summing up in plenum. The aim of this exercise is to increase awareness of what changes life with a newborn has on a daily routine, how much time is spent on breastfeeding, etc. [14]. Participants are also asked to consider activities that give them energy and pleasure (e.g. playing football, going out with friends, reading a book), and how they might incorporate some of these activities in their new daily lives [14]. Next participants are encouraged to reflect upon how they normally handle worries, and to discuss this topic with their partner. Finally common emotional reactions and postnatal depressive symptomatology is covered. The importance of being open about these emotions and supportive of one another is stressed, as is the importance of seeking help when deemed necessary.

Session 4 (approximately 5 weeks post-partum):

- Birth experiences.
- Mood swings and postnatal depressive symptomatology.
- The first time at home with a newborn.
- Couplehood.

This session is carried out by a midwife as well as a healthcare visitor. The aim is for the newborn parents to share birth experiences, and their experiences in their new roles as parents so far. The topic of common emotions and postnatal depressive symptomatology is revisited. Next groups of four are asked to discuss how being a parent is different to what they expected, which challenges they consider to be the greatest, how they cope with/handle these challenges, and what have been the greatest joys. The parents are able to ask the midwife and the healthcare visitor practical questions during the break. After the break, using the communication framework, couples are asked to discuss what the best change has been in their partner after becoming a parent, what has worked really well in the relationship, and in sharing household tasks, and what could make it even better. Finally the healthcare visitor talks about sex (including contraception) and intimacy after becoming parents.

Pilot study

The feasibility and face validity of the programme has been pilot- tested among 35 couples by qualitative interviews, and observation of participants and facilitators (unpublished data).

Ethics–risk/benefits

The trial is approved by and registered with the regional ethics committee, and will be carried out in accordance with the Declaration of Helsinki in its latest form as well as national laws and regulations.

There are no known risks of participating in the trial. We assume that participants in the experimental group will benefit from more in depth antenatal classes in small groups, however, we cannot rule out the possibility that the experimental group may experience an initial increase in worries about issues related to birth and parenting. Participants randomised to the control group may experience some disappointment. To date there is no conclusive evidence as to which form and format an antenatal programme should have. We therefore consider it ethically justifiable that the control group will receive standard care.

Participants are free to attend concomitant antenatal and postnatal services and parent groups. Participants will be able to withdraw from the trial at any time. Women who have a miscarriage or a stillborn child will not be continued in the trial.

Data collection

All Danes have a unique personal identification number (CPR-number) which identifies sex, date and year of birth and allows for register linkage with all population-based registers in Denmark. Data will be collected by the hospital obstetric database, the national registers, and web-based questionnaires from both parents at: baseline, i.e. time point 0 (tp 0) (at approximately 18 weeks of gestation); at 37 weeks of gestation (tp 1); at 9 weeks after expected due date (tp 2); at 6 months after expected due date (tp 3); and 1 year after expected due date (tp 4). Participants will be contacted via e-mail, and a reminder will be sent by e-mail after a week. After 2 weeks phone numbers will be sought on participants who have still not responded, and where possible they will be contacted by phone.

Blinding

It is not possible to blind the participants and the personnel in the trial. However, blinding in all other aspects of the trial will be maintained: blinded data collection on outcomes from national registers; the statistical analyses will be conducted with the two intervention groups coded as, e.g. A and B; and two conclusions will be drawn by the Steering Committee, one assuming A is the experimental group and B is the control group, and one conclusion assuming the opposite. After this the blinding will be broken.

Outcome measures

As blinding of participants and midwives and health visitors is not feasible in this trial it is desirable to specify at least one objectively assessed outcome to reduce the risk of bias, even if the outcome of most interest is subjective [17].

The primary outcome is use of epidural analgesia during labour, using data from the hospital obstetric database (as proxy variable for coping and fear of childbirth). Findings suggest that women who receive epidural analgesia experience more fear but not more pain, before the administration of epidural analgesia [18]. Structured antenatal education may improve women's ability to cope during labour resulting in lower epidural rates [8].

Secondary outcomes are: stress measured by The Swedish Parenthood Stress Questionnaire (SPSQ) [19] – questionnaire data (tp 2,3,4), and The Perceived Stress Scale (PSS) [20] – questionnaire data (tp 0,1,2,3,4). Parenting alliance – The Parenting Alliance Measure [21] – questionnaire data (tp 2,3,4)

Explorative outcomes: antenatal and postnatal depressive symptomatology and anxiety – Edinburgh Postnatal Depression Scale (tp 0,1,2), The Major Depression Inventory (MDI) (tp 3), The Hopkins Symptom Check List (SCL-25) first 10 items (anxiety score,

SCL-anxiety)(12)(tp 1,2) – questionnaire data. Breastfeeding – questionnaire data (tp 0,1,2,3). Use of healthcare services, i.e.: for the parents obstetric intervention, i.e. augmentation of labour, vacuum extraction, caesarean section – data from the hospital obstetric database (tp 2), and contact to healthcare professionals for depressive symptomatology and unscheduled postnatal visits–questionnaire data (tp 2,3). For the child, i.e. neonatal readmissions to hospital, contacts to accident and emergency departments (A and E), General Practitioner (GP) and doctor on call during the child's first year of life (composite measure). Use of the regional emergency phone line – data from the national registers (tp 4). Family medicine use and smoking – questionnaire data (tp 2,3) and register data (tp 4). Satisfaction with relationship and family break-ups – questionnaire data (tp 0,1,2,3), and data from the national registers on divorce and break-ups (tp 4). Mental well-being The Warwich-Edinburgh Mental Well-being scale (tp 1,2,3) [13].

Intermediate outcomes: parenting resources: Self-efficacy in relation to: (1) birth (tp 1), (2) discharge (tp 1), (3) parenting (tp 1,2), (4) breastfeeding (tp 0,1); couple communication (tp 0,1,2,3); social support/network (tp 0,1,2,3).

Demographic variables and individual characteristics: education, occupational social class, marital status, cohabiting status and number of children living in the household, sense of coherence, self-rated health, subjective health complaints, and long term illness.

The questionnaires include between 70 and 190 items, and take between 10 and 25 min to complete.

Process evaluation will be conducted with a mixed methods approach using quantitative questionnaire data and qualitative interviews with participants and service providers. We will examine programme fidelity, e.g. whether the protocol is followed in programme delivery, and how much of the intended programme the participants receive (dose), as well as programme reach, e.g. what proportions of the intended groups are participating in the programme, as these factors have an impact on the effect of the intervention [22]. Participants will be asked to fill in a web-based questionnaire on-site at the end of each session. The questionnaire will highlight whether the intended educational subjects of the session have been covered as well as the extent to which the participants found the information given useful. Group facilitators will be asked to fill in a similar questionnaire with the opportunity to explain why certain topics may have been omitted.

Furthermore, participant observation will be carried out during random sessions. Using in-depth interviews qualitative data will be collected from a purposive sample of participants to gain understanding of their perceptions and experiences of the NEWBORN course using interview schedules with topic guides.

Use of additional antenatal and postnatal services will be investigated by questions specifically developed for this purpose, and analysis adjusted for concomitant service use will be performed.

Cost-effectiveness analysis: the incremental societal cost of the intervention will be computed and compared to the measured outcomes. Direct health care costs and productivity costs in terms of labour market participation and short term absence will be calculated.

Statistical plan and data analysis

Sample size

We are planning a trial of experimental and control participants with one control participant per experimental participant. 2011 data from the HH Obstetric Database [23] indicate that use of epidural analgesia among pregnant women is 31%. If the true use of epidural analgesia for experimental participants is 25% (risk reduction of 19%), we will need to include 1175 participants in the experimental group and 1175 control participants to be able to reject the null hypothesis that the epidural use for experimental and control participants is equal with a probability (power) of 90%. The type I error probability associated with this test of this null hypothesis is 5%.

For the three secondary outcomes, we have estimated the power of 98% or more (Table 1).

Statistical methods

Reporting will follow the guidelines of the CONSORT-statement. Statistical analyses will be intention to treat and per protocol. The level of significance is set to 5% and power to 90%.

The analysis of the primary binary outcome will be done using the generalised linear mixed model with distribution = binomial, link = logit and 'experimental antenatal group' as a random factor. The 'antenatal subgroup' comprise the groups of 4–6 couples for the experimental participants and the collective group of control participants. The analysis will be adjusted by the protocol specified stratification variables.

If the percent missing cases >5%, the results of this analysis will be subjected to a 'worst case' and a 'best case' scenario analysis of the potential impact of missing values. Assume a beneficial effect (less use of epidural analgesia) is noted in one group (group A) as compared to the other group (group B). A worst case scenario will then be constructed where missing values in group A are imputed by a "yes" to use of epidural analgesia and missing values in group B are imputed by a "no" to use of epidural analgesia. A corresponding best case scenario will also be constructed and the result under both scenarios will be computed.

Analyses (adjusted by baseline value and protocol specified stratification variables) of perceived stress at 37 week gestation and for each of the other three secondary outcomes of the area under the curve (AUC) from 9 weeks to 1 year after due date will be done. The linear mixed model with the intervention indicator as a fixed effect and group as a random effect will be used in the analyses.

If the percent missing cases of an outcome is >5% and p of Little's test (1) <5%, a number of datasets with observed values and predicted unobserved values necessary to produce an efficiency >99% will be produced using multiple imputations (SPSS version 17 or later) (2). The primary analysis will then be that based on these data sets.

Multiplicity will be dealt with as follows: the primary outcome will be tested at the 5% level. If $p > 5\%$, the remaining null hypotheses will be accepted without test. If not, the p values of the

Table 1
Power estimations for the secondary outcomes in a trial with 2350 participants.

	Number of participants	Minimal relevant difference	Standard deviation	Type 1 error risk (%)	Power (%)*
Perceived stress scale	2350	1 Point [15]	6 Points [15]	5	98.1
Swedish Parenthood Stress Questionnaire	2350	0.1 Point [19]	0.5 Points [19]	5	99.8
Parenting Alliance Measure	2350	4 Points [21]	20 Points [21]	5	99.8

* Power estimations conducted with the programme: PS Power and Sample Size Calculations version 3.0.14 [26].

remaining four tests will be adjusted using Hommel's procedure [24]. In all events all observed p values will be reported.

Discussion

Antenatal education classes are offered to prospective parents in most countries in the Western part of the world. However, there is very limited knowledge on the effect of, as well as the content and form of antenatal education.

To our knowledge, the NEWBORN trial is the largest randomised trial to date. We will minimise the risk of bias in all important domains [25]. Although it is impossible to blind participants and investigators, we will be able to blind all other aspects of the trial. Due to the comprehensiveness of Danish registers, we will obtain blinded and objective assessment of the primary outcome.

The trial recruits participants from a single hospital in Denmark, which may reduce the external validity of findings. However, the experimental intervention is delivered by 20 different midwifes and 8 different health-visitors in 3 different local sites, which in turn will increase the generalisability. Further, the trial has very wide eligibility criteria, leaving potential findings applicable to the entire Danish population.

Previous trials and studies have mainly focused on the mother's transition into motherhood. In the NEWBORN trial we will have a strong focus on the father and the couple as a whole. This will bring valuable new knowledge to an area with limited knowledge.

It may be seen as a limitation that the experimental and the control group differ in more than one respect. The size of the groups differ and the type of teaching. They also differ in terms of the actual material presented. The study provides a comparison between the two approaches. But safe inferences pertaining to the causes of a difference between the two approaches regarding type of teaching, content of teaching and duration of teaching cannot be made.

To date, we have only planned follow-up till 1 year after due date. This leaves several limitations regarding the assessment of participant-relevant outcomes, such as the child's thriving as it grows up, the number of families that experiences divorces and break-ups, and child's use of the health-care system in both the short and the long run. We assess these outcomes in the NEWBORN trial, but we range them in the outcome hierarchy as 'exploratory'. This is done, as (1) we have very limited knowledge of the potential effect of antenatal education on these outcomes, and we have therefore not been able to perform power estimations as we have for the secondary outcomes. (2) Due to logistical and financial constraints. If additional funding can be obtained, data on all individuals can be sought in the national registers and long-term follow-up is possible.

Results from the NEWBORN trial will form a much-needed base for decision-makers regarding the form and content of antenatal education.

Trial status

Recruitment of participants started November 2012. No interim analyses have been conducted. Data collection on the primary outcome is expected to be complete medio 2014. Full data collection is expected to be complete medio 2015.

Conflict of interest

The authors declare that they have no conflict of interests.

Author's contributions

Appendix A shows how authors have contributed to the trial. All authors have read and approved the final manuscript.

Finance and insurance

The Danish Cancer Society has funded the project with 5.2 MDKK (approximately 900.000 USD) but have not been involved in any aspects related to the study of the decision to publish.

References

[1] Gagnon AJ, Sandall J. Individual or group antenatal education for childbirth or parenthood, or both. Cochrane Database Syst Rev 2007;3:CD002869.
[2] Hvidovre Hospital. Description of the content of Danish antenatal classes over time. Hvidovre Hospital; 2012.
[3] Hillerød Hospital. Description of the content of Danish antenatal classes over time. Hillerød Hospital; 2012.
[4] Rigshospitalet. Description of the content of Danish antenatal classes over time. Rigshospitalet; 2012.
[5] National Board of Health D. Status on regional and municipal offers in relation to pregnancy and delivery with focus on delivery and early discharge (In Danish: Status for regionernes og kommunernes tilbud til gravide og fødende med fokus på ambulante fødsler og tidlige udskrivelser). Copenhagen: 2011. Report No.: j.nr. 7–205-03-31/1/chb.
[6] Cowan C, Cowan P. When partners become parents: the big life change for couples. Mahwah, NJ: Lawrence Erlbaum Associates Publishers; 2000. 2000.
[7] Illeris K, Jarvis P, Kegan R, Engeström Y, Elkjaer B, Mezirow J, et al. Contemporary theories of learning: learning theorists...in their own words. New York: Routledge; 2009.
[8] Maimburg RD, Vaeth M, Durr J, Hvidman L, Olsen J. Randomised trial of structured antenatal training sessions to improve the birth process. BJOG 2010;117(8):921–8.
[9] Svensson J, Barclay L, Cooke M. Randomised-controlled trial of two antenatal education programmes. Midwifery 2009;25(2):114–25.
[10] Barlow J, Coren E. Parent-training programmes for improving maternal psychosocial health. Cochrane Database Syst Rev 2004;1:CD002020.
[11] Barlow J, Parsons J. Group-based parent-training programmes for improving emotional and behavioural adjustment in 0–3 year old children. Cochrane Database Syst Rev 2003;1:CD003680.
[12] Sanders MR. Triple P-positive parenting program as a public health approach to strengthening parenting. J Fam Psychol 2008;22(4):506–17.
[13] Freedman CM, Low SM, Markman HJ, Stanley SM. Equipping couples with the tools to cope with predictable and unpredictable crisis events: the PREP program. Int J Emerg Mental Health 2002;4(1):49–55.
[14] Milgrom J, Schembri C, Ericksen J, Ross J, Gemmill AW. Towards parenthood: an antenatal intervention to reduce depression, anxiety and parenting difficulties. J Affect Disord 2011;130(3):385–94.
[15] Bandura A. Self-efficacy: the exercise of control. New York: W. H. Freeman; 1997.
[16] Sundhedsoplysning Kf. Kort og godt om amning. 11th ed. Copenhagen, Denmark: Komiteen for Sundhedsoplysning; 2013.
[17] Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. BMJ 2008;336(7644):601–5.
[18] Alehagen S, Wijma K, Wijma B. Fear during labor. Acta Obstet Gynecol Scand 2001;80(4):315–20.
[19] Ostberg M, Hagekull B, Wettergren S. A measure of parental stress in mothers with small children: dimensionality, stability and validity. Scand J Psychol 1997;38(3):199–208.
[20] Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. J Health Soc Behav 1983;24(4):385–96.
[21] Abidin RR, Brunner JF. Development of A parenting alliance inventory. J Clin Child Psychol 1995;24(1):31–40.
[22] Bartholomew LK, Parcel GS, Kok G, Gottlieb NH, Fernández ME. Planning health promotion programs – an intervention mapping approach. 3rd ed. Francisco: Jossey-Bass; 2011.
[23] Hvidovre Hospital. Data from the obstetric database. Hvidovre Hospital; 2012.
[24] Konietschke F. Multiple testing problems in pharmaceutical statistics. AU – A. Dmitrienko, AU – A. C. Tamhane, and AU – F. Bretz (2010). Boca Raton: PB – Chapman & Hall/CRC Biostatistics Series. ISBN: 978-1-584 88984-7. Biom J 2012;54(4):579–80.
[25] Higgins JPT, Green S. Cochrane handbook for systematic reviews of interventions version 5.1.0.; 2011.
[26] Dupont WD, Plummer Jr WD. Power and sample size calculations. A review and computer program. Control Clin Trials 1990;11(2):116–28.

Paper IV

Clinical Epidemiology

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Validity of a hospital-based obstetric register using medical records as reference

This article was published in the following Dove Press journal:
Clinical Epidemiology
26 November 2015
Number of times this article has been viewed

Carina Sjöberg Brixval[1]
Lau Caspar Thygesen[1]
Nanna Roed Johansen[2]
Christina Rørbye[3]
Tom Weber[3]
Pernille Due[1]
Vibeke Koushede[1]

[1]National Institute of Public Health, University of Southern Denmark, Copenhagen, Denmark; [2]Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark; [3]Copenhagen University Hospital, Department of Obstetrics and Gynecology, Hvidovre Hospital, Hvidovre, Denmark

Background: Data from hospital-based registers and medical records offer valuable sources of information for clinical and epidemiological research purposes. However, conducting high-quality epidemiological research requires valid and complete data sources.

Objective: To assess completeness and validity of a hospital-based clinical register – the Obstetric Database – using a national register and medical records as references.

Methods: We assessed completeness of a hospital-based clinical register – the Obstetric Database – by linking data from all women registered in the Obstetric Database as having given birth in 2013 to the National Patient Register with coverage of all births in 2013. Validity of eleven selected indicators from the Obstetric Database was assessed using medical records as a golden standard. Using a random sample of 250 medical records, we calculated proportion of agreement, sensitivity, specificity, and positive and negative predictive values for each indicator. Two assessors independently reviewed medical records and inter-rater reliability was calculated as proportion of agreement and Cohen's κ coefficient.

Results: We found 100% completeness of the Obstetric Database when compared to the Danish National Patient Register. Except for one delivery all 6,717 deliveries were present in both registers. Proportion of agreement between the Obstetric Database and medical records ranged from 91.1% to 99.6% for the eleven indicators. The validity measures ranged from 0.70 to 1.00 indicating high validity of the Obstetric Database. κ coefficients from the inter-rater reliability ranged from 0.71 to 1.00.

Conclusion: Completeness and validity of the Obstetric Database were found acceptable when using the National Patient Register and medical records as golden standards. The Obstetric Database therefore offers a valuable source for examining clinical, administrative, and research questions.

Keywords: obstetric register, register-based, hospital register, validity, completeness

Introduction

In Denmark, approximately 60,000 children are born each year. During the past years, the proportion of interventions in the birth process has increased, eg, the rate of epidural analgesia has increased from 18% in 2006 to 24% in 2013.[1] Monitoring prevalence and time trends in health outcomes and medical procedures requires valid and complete data sources. All residents in Denmark are included in the Danish health registers and accurate linkages are possible due to the unique personal identification number[2] making Denmark a suitable setting for register-based research.

The advantages of register-based research is the representativeness of the study population, that risk of diagnostic process is not affected by the research question, and that data already exist which minimize time consumption and costs. A disadvantage

Correspondence: Carina Sjöberg Brixval
National Institute of Public Health, University of Southern Denmark, Øster Farigmagsgade 5A, 1353 Copenhagen, Denmark
Tel +45 6550 7845
Email cabr@niph.dk

of using register-based data for research purposes is that data collection and data quality are not under control of the researcher and information on more specific and detailed treatment procedures of clinical interest present in hospital-based clinical registers is often not included in national registers.[3,4] Therefore, clinical registers are often necessary when conducting clinical epidemiological research.

Validation of register-based data is necessary to ensure the quality of the data. A well-conducted validation study includes sample size calculations[5] and comparison of information in the given register with information in other registers as well as primary data sources, eg, medical records.[6] The use of two assessors to extract data when validating information from registers against medical records reduces the risk of registration error. An examination of the inter-rater reliability will also give an indication of how difficult a given indicator in the medical record is to assess.

The Obstetric Database at Hvidovre Hospital has been used for internal monitoring of prevalence of obstetric interventions as well as research and high validity is therefore essential. The Obstetric Database has not been validated previously and the aim of this study was therefore to assess completeness and validity of the Obstetric Database using the Danish National Patient Register and medical records as references. Furthermore, the detailed description of the validation process in this paper may be helpful to fellow researchers or clinicians wanting to examine completeness and validity of a clinical register.

Methods

We assessed the completeness and validity of the Obstetric Database using three data sources; the National Patient Register, the Obstetric Database, and medical records. Assessment of completeness of the Obstetric Database was performed by comparing information on the presence of the unique personal identification number in the National Patient Register and the Obstetric Database. The validity of the Obstetric Database was assessed by comparing information on the presence of selected indicators in the Obstetric Database and medical records.

Registers

The Obstetric Database is a hospital-based clinical register comprising initial obstetric and neonatal data, eg, Apgar score, obstetric interventions and outcomes from all deliveries at the three largest delivery wards (Herlev Hospital, Rigshospitalet, and Hvidovre Hospital) in the Capital Region of Denmark. At Hvidovre Hospital, the database was established in 1996. We selected Hvidovre Hospital to assess validity of the Obstetric Database as this is the largest labor ward in Denmark with more than 6,500 deliveries each year. During and after labor, midwives register the obstetric and neonatal baseline data and interventions in the database by ticking an electronic list. Postpartum, a specialist doctor or senior midwife goes through every file and adds left out information and supplies specialist diagnoses based on information from the medical records.

The Danish National Patient Register was established in 1977 and records in- and outpatient contacts from all hospitals in Denmark. The register contains information on the dates of admission and discharge, and information on diagnoses and major clinical procedures performed at hospitals. The purpose of the National Patient Register is among others to provide information for the production of statistical data and serve as the basis for the payment of hospitals via the Diagnostic Related Group system.[7]

Completeness assessment

We assessed the completeness of the Obstetric Database by using the National Patient Register as a gold standard. Reporting to the National Patient Register is compulsory to obtain reimbursement from health authorities for patient contacts and we therefore anticipate this register to be complete. Completeness of the Obstetric Database was defined as the percentage of deliveries in the Obstetric Database that were also registered in the National Patient Register. Registrations to the National Patient Register are performed by the secretaries at the Department of Obstetrics and Gynecology at Hvidovre Hospital based on information from the Obstetric Database. For the assessment of completeness of the Obstetric Database, we selected data on all deliveries planned to be performed at Hvidovre Hospital in 2013. In the Obstetric Database, all planned deliveries at a given hospital are registered under that hospital regardless of whether or not they actually ended up being carried out there. Stillbirths were also included. Twin- and triplet births counted as one delivery. All deliveries at Hvidovre Hospital in the Obstetric Database and the National Patient Register were linked by the unique personal identification number.

Validity of selected indicators

The validity of the Obstetric Database was assessed using medical records as gold standard. For the purpose of this study, the following eleven indicators were considered of most importance to be validated: use of oxytocin due to dystocia, epidural analgesia, vacuum extraction, emergency and elective cesarean

delivery, perineal tear – degree 3 and 4, test for gestational diabetes, scalp blood pH, arterial pH from the umbilical cord, and Apgar score after 5 minutes. We assessed the presence of the indicators in the Obstetric Database and medical records. Three of the indicators are continuous measures (scalp blood pH, arterial pH from the umbilical cord, and Apgar score after 5 minutes). We did not assess the agreement of the values of these indicators between the Obstetric Database and the medical records. Some of the indicators are obstetric interventions such as administration of oxytocin and cesarean delivery only performed among patients with complications; others are routine measurements such as arterial pH from the umbilical cord and Apgar score. Tests for gestational diabetes are performed only among pregnant women with an increased risk of diabetes.

Sample size calculations were based on the primary measure of validity; positive predictive value. We expected a valid registration of 95% between the Obstetric Database and medical records and wanted to estimate this with a confidence interval of 92%–98%. In order to fulfill this, a sample of 203 deliveries was required. To take into account that data on certain outcomes might be incomplete due to, eg, stillbirth, and further that agreement between some outcomes could potentially be lower than the expected 95%, a random sample of 250 deliveries in 2013 was selected from the Obstetric Database and the corresponding electronic medical records were retrieved. We excluded two deliveries not performed at Hvidovre Hospital due to transfers to other hospitals during labor and one delivery due to missing data on all indicators for unknown reasons leaving 247 deliveries for validation. These exclusions were necessary due to lack of information in the medical records for these women and their children. For the assessment of validity of arterial pH from the umbilical cord and Apgar score after 5 minutes, we additionally excluded three records due to stillbirths leaving 244 deliveries for validation of these indicators.

Blinded for information in the Obstetric Database, two of the authors (NRJ and CSB) independently assessed the medical records for registration of the aforementioned indicators. In cases where Apgar score or arterial pH from the umbilical cord was not present in the mother's medical record, these indicators were assessed from the child's medical record. Assessment of performance of test for gestational diabetes was estimated from an overview of blood test results found in a separate section of the medical record. Data from each assessor was entered into separate Excel sheets. In cases of doubt as to whether an obstetric intervention had been performed, the authors consulted two skilled obstetricians (CR and TW) for clarification independently of one another. Next, the datasets from the two assessors were compared and any disagreements were solved by an obstetrician (CR) who was blind to the assessor. The final dataset was linked to data in the Obstetric Database.

Statistical analyses

Sensitivity, specificity, positive and negative predictive values, as well as proportion of agreement were calculated for each of the eleven indicators. We calculated exact 95% binomial proportion confidence intervals. Definitions of the validity measures are given in Table 1.

Inter-rater reliability between the two assessors after consultation with the obstetricians was calculated as proportion of agreement as well as Cohen's κ coefficient for each of the indicators. We used the Landis and Koch's scale[8] to categorize strength of agreement from the κ coefficients.

All statistical analyses were performed using SAS version 9.3, SAS Institute Inc., Cary, NC, USA.

Ethical issues

This study fulfills all Danish ethical standards and was approved by the Danish Data Protection Agency (No 2014-54-0714) and by the Department of Obstetrics and Gynecology at Hvidovre Hospital.

Results

Completeness

In the Obstetric Database, 6,718 deliveries were registered in 2013, whereas 6,717 deliveries were registered in the Danish National Patient Register. When linking data from the Obstetric Database and the National Patient Register,

Table 1 Definition of measures of validity

Obstetric Database	Medical record (gold standard)		Total
	Present	Absent	
Present	a	b	a + b
Absent	c	d	c + d
Total	a + c	b + d	

Notes: The sensitivity is the proportion of patients with registration of the indicator according to both medical records and the Obstetric Database (a), compared to all patients with the indicator according to medical records (a + c) = True positive (a)/(True positive [a] + false negative [c]). The specificity is the proportion of patients without registration of the indicator according to both medical records and the Obstetric Database (d), compared to all patients without the indicator according to medical records (b + d) = True negative (d)/(True negative [d] + false positive [b]). The positive predictive value is the proportion of patients with registration of the indicator according to both medical records and the Obstetric Database (a), compared to all patients with the indicator according to the Obstetric Database (a + b) = True positive (a)/(True positive [a] + false positive [b]). The negative predictive value is the proportion of patients without registration of the indicator according to both medical records and the Obstetric Database (d), compared to all patients without the indicator according to the Obstetric Database (c + d) = True negative (d)/(True negative [d] + false negative [c]).

Table 2 Number of registrations in the Obstetric Database and in medical records and the proportion of agreement (%) for each indicator

Indicator	In the Obstetric Database		Not in the Obstetric Database		Proportion of agreement, %, (95% confidence interval)
	In medical records	Not in medical records	In medical records	Not in medical records	
Oxytocin due to dystocia	46	20	2	179	91.1 (86.8–94.3)
Epidural analgesia	68	0	3	176	98.8 (96.5–99.8)
Vacuum extraction	21	0	2	224	99.2 (97.1–99.9)
Emergency cesarean delivery	35	0	1	211	99.6 (97.8–100.0)
Elective cesarean delivery	23	1	1	222	99.2 (97.1–100.0)
Perineal tear degree 3	8	1	0	238	99.6 (97.8–100.0)
Perineal tear degree 4	0	0	0	247	100.0 (100.0–100.0)
Scalp blood pH	69	1	0	177	99.6 (97.8–100.0)
Arterial pH from the umbilical cord	221	3	3	17	97.9 (94.7–99.1)
Apgar score after 5 minutes	241	2	1	0	98.8 (96.5–99.8)
Test for gestational diabetes	84	6	9	148	93.9 (90.2–96.6)

6,717 deliveries were present in both data sources. No deliveries present in the National Patient Register were missing in the Obstetric Database and only one delivery was present in the Obstetric Database but not in the National Patient Register, indicating almost exact agreement (rounded to 100%) between the two registers.

Validity of indicators

The proportion of agreement between the Obstetric Database and medical records was high for most indicators (Table 2). For nine of the indicators, the proportion of agreement was 97.9% or above. Oxytocin due to dystocia (91.1%) and test for gestational diabetes (93.9%) had lower proportions of agreement.

Sensitivity for all indicators was high and ranged from 0.90 (test for gestational diabetes) to 1.00 (perineal tear degree 3, scalp blood pH, and Apgar score after 5 minutes) (Table 3). Also, specificity was high and ranged from 0.85 (arterial pH from the umbilical cord) to 1.00 (epidural analgesia, vacuum extraction, emergency and elective cesarean delivery, and perineal tear degree 3).

The predictive values were generally high. Except for oxytocin due to dystocia, the positive predictive values ranged from 0.89 (perineal tear degree 3) to 1.00 (epidural analgesia, vacuum extraction, and emergency cesarean delivery). However, the results revealed false positive registrations of the indicator oxytocin due to dystocia in the Obstetric Database resulting in a positive predictive value of 0.70.

Negative predictive values ranged from 0.85 (arterial pH from the umbilical cord) to 1.00 (emergency and elective cesarean delivery, perineal tear degree 3, and scalp blood pH).

Inter-rater reliability

Proportion of agreement between the two assessors ranged from 94.3% (oxytocin due to dystocia) to perfect agreement

Table 3 Sensitivity, specificity, positive, and negative predictive values (95% confidence interval) for eleven indicators in the Obstetric Database

Indicator	Sensitivity	Specificity	Positive predictive value	Negative predictive value
Oxytocin due to dystocia	0.96 (0.86–0.99)	0.90 (0.85–0.94)	0.70 (0.57–0.80)	0.99 (0.96–1.00)
Epidural analgesia	0.96 (0.88–0.99)	1.00 (0.98–1.00)	1.00 (0.95–1.00)	0.98 (0.95–1.00)
Vacuum extraction	0.91 (0.72–0.99)	1.00 (0.98–1.00)	1.00 (0.84–1.00)	0.99 (0.97–1.00)
Emergency cesarean delivery	0.97 (0.85–1.00)	1.00 (0.98–1.00)	1.00 (0.90–1.00)	1.00 (0.97–1.00)
Elective cesarean delivery	0.96 (0.79–1.00)	1.00 (0.98–1.00)	0.96 (0.79–1.00)	1.00 (0.98–1.00)
Perineal tear degree 3	1.00 (0.63–1.00)	1.00 (0.98–1.00)	0.89 (0.52–1.00)	1.00 (0.98–1.00)
Perineal tear degree 4[a]	–	–	–	–
Scalp blood pH	1.00 (0.95–1.00)	0.99 (0.97–1.00)	0.99 (0.92–1.00)	1.00 (0.98–1.00)
Arterial pH from the umbilical cord	0.99 (0.96–1.00)	0.85 (0.62–0.97)	0.99 (0.96–1.00)	0.85 (0.62–0.97)
Apgar score after 5 minutes[b]	1.00 (0.98–1.00)	0	0.99 (0.97–1.00)	0
Test for gestational diabetes	0.90 (0.82–0.95)	0.96 (0.92–0.99)	0.93 (0.86–0.98)	0.94 (0.89–0.97)

Notes: [a]No perineal tear degree 4 was registered in either Obstetric Database or medical records. Therefore, statistics are not presented for this indicator; [b]specificity and negative predictive value equals 0 as Apgar score after 5 minutes was always registered in either Obstetric Database or the medical record or both (no true negative).

Table 4 Prevalence of indicators in medical records by each assessor, inter-rater agreement (%) and κ coefficients (95% confidence interval) for each indicator

Indicator	Assessor 1	Assessor 2	% agreement	κ coefficient
Oxytocin due to dystocia	51	57	94.3	0.83 (0.75–0.92)
Epidural analgesia	71	71	100	1.00 (1.00–1.00)
Vacuum extraction	23	24	99.6	0.98 (0.93–1.00)
Emergency cesarean delivery	34	36	99.2	0.97 (0.92–1.00)
Elective cesarean delivery	26	24	99.2	0.96 (0.89–1.00)
Perineal tear degree 3	7	7	98.4	0.71 (0.43–0.98)
Perineal tear degree 4[a]	0	1	99.6	–
Scalp blood pH	68	69	99.6	0.99 (0.97–1.00)
Arterial pH from the umbilical cord	227	228	97.1	0.77 (0.61–0.94)
Apgar score after 5 minutes	242	241	99.6	0.80 (0.41–1.00)
Test for gestational diabetes	85	93	95.1	0.89 (0.84–0.95)

Note: [a]It was not possible to calculate the κ coefficient for perineal tear degree 4 due to no registrations for assessor 1.

of 100% for epidural analgesia (Table 4). κ coefficients ranged from 0.71 for perineal tear degree 3 to a perfect agreement of 1.00 for epidural analgesia. Using perineal tear degree 3 as an example: although both assessors noted seven events only five of these seven events were the same, resulting in an agreement of 98.4%.

Discussion

We examined completeness and validity of a hospital-based clinical register at the largest birth site in Denmark.

Completeness was assessed by comparing data from all women registered in the Obstetric Database as having given birth in 2013 and linking to the National Patient Register which was considered a gold standard. We found that all deliveries registered in the National Patient Register were also registered in the Obstetric Database, giving a completeness of 100%. One delivery was not registered in the National Patient Register for unknown reason. Registrations to the National Patient Register are based on information from the Obstetric Database. Reporting to the National Patient Register is compulsory to obtain reimbursement from health authorities for patient contacts and we therefore considered this register as a gold standard. This study supports that the Obstetric Database is used very actively in the clinical practice and that no deliveries therefore are missing.

We used medical records as gold standard when assessing validity of the Obstetric Database and found that sensitivity, specificity, and predictive values generally were high for the selected eleven indicators indicating high validity of the database.

For all indicators sensitivity and specificity was high (≥0.91 and ≥0.85 respectively). This implies that the Obstetric Database has high validity regarding registration from the medical record. Also, the predictive values were generally very high indicating a high probability that the registrations in the Obstetric Database are correct.

A previous systematic review of perinatal validation studies have shown that indicators related to type of delivery and perineal tear are well reported with high sensitivities and positive predictive values, whereas induction and augmentation of labor have higher degrees of underreporting.[9] The results from the present study are thus in accordance with former validity studies in the obstetric field.

Although the proportion of true positive results (the positive predictive value) in the Obstetric Database was high for almost all indicators, 20 cases of oxytocin due to dystocia were registered in the Obstetric Database but not in the medical records. In all these instances, use of oxytocin occurred as part of induction of labor according to the medical records. According to the Danish guidelines for registration of obstetric interventions, oxytocin administration should only be coded as induction if it is used as the first procedure for induction of labor. If oxytocin is administered as a secondary induction procedure it is coded as "due to dystocia".[10] In two of the aforementioned 20 cases, oxytocin was registered in the Obstetric Database as induction, while ten others were registered as induced with Misoprostol before treatment with oxytocin. The remaining eight were not registered as induced. During recent years, the registration practice has changed. Previously, indications for oxytocin administration were registered separately for induction of labor and for dystocia.[10] While the former registration practice provided an opportunity for assessing oxytocin due to dystocia and oxytocin as induction procedure separately, this is no longer possible due to the current registration practice. This implies a potential risk of misinterpretation of data if one wishes to study oxytocin due to dystocia.

Validity was also lower for test for gestational diabetes. Six women were registered with the test in the Obstetric Database but not in the medical record. If the person entering data into the Obstetric Database did not check whether the test was actually performed, they may have falsely registered tests based on recorded indication alone. The nine tests for gestational diabetes registered in the medical record but not in the Obstetric Database have most likely been overlooked by the person entering data into the Obstetric Database. This could be due to complex registration systems, ie, that information has to be found in separate sections of the medical records.

The κ coefficients for the inter-rater reliability was above 0.70 for all the indicators, and is therefore considered either "substantial" or "almost perfect" according to the Landis and Koch categorization.[8] In cases of disagreements between the two assessors the decision was made by a skilled obstetrician. We therefore consider the reliability of the review of the medical records to be adequate and the medical records to be valid as gold standard. The κ coefficients were lower for the indicators that also had low predictive values (oxytocin due to dystocia, perineal tear degree 3, and arterial pH from the umbilical cord) indicating that these indicators were generally more difficult to assess.

Strengths of this study include the use of the National Patient Register with national coverage as well as medical records as gold standards. Further, the extensive review of medical records was performed by two independent assessors and approved by two independent clinical experts. The high agreement between the two assessor's registrations confirms that use of the medical records as gold standards was appropriate. The random sample selected among women giving birth at Hvidovre Hospital makes these results generalizable to all deliveries at Hvidovre Hospital in this period. The results might not be generalizable to other clinical databases at other hospitals as registration practices might vary between hospitals. However, the registration guidelines for the obstetric coding apply throughout the entire country which suggests that the results may be generalizable to other clinical databases.

We assessed whether the indicators were present in the medical records and the Obstetric Database. The accuracy of the values of scalp blood pH, arterial pH from the umbilical cord, and Apgar score was not assessed. Therefore, further validation of the accuracy of these indicators will be desirable before using them for research or administrative purposes.

Both the issue of using the Danish National Patient Register as gold standard and the reduced validity of a few of the indicators stress the importance of careful consideration and evaluation of the completeness and validity of the different components of registers.

In conclusion, completeness and validity of the selected indicators in the Obstetric Database are high. With data being valid and the database complete, the Obstetric Database offers a valuable source for monitoring prevalence of obstetric interventions and outcomes as well as obstetrical research studies. However, when monitoring use of oxytocin due to dystocia, care should be taken as the code for this obstetric intervention might also cover oxytocin used as part of induction of labor.

The detailed description of the validation process may be helpful to fellow researchers or clinicians wanting to examine completeness and validity of a clinical register.

Acknowledgments

This study was conducted as part of a PhD study funded by the Danish Cancer Society. The Danish Cancer Society does not have any influence in study design, in the collection, analysis, and interpretation of data, in the writing of the manuscript, or in the decision to submit this manuscript.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Statens Serum Institut. Sygehusfødte og komplikationer [Deliveries at hospitals and complications] [database on the Internet]; 2014. Available from: http://www.esundhed.dk/sundhedsregistre/MFR/Sider/MFR03.aspx. Accessed January 22, 2015. Danish.
2. Thygesen LC, Daasnes C, Thaulow I, Bronnum-Hansen H. Introduction to Danish (nationwide) registers on health and social issues: structure, access, legislation, and archiving. *Scand J Public Health*. 2011;39(7 Suppl):12–16.
3. Sorensen HT, Sabroe S, Olsen J. A framework for evaluation of secondary data sources for epidemiological research. *Int J Epidemiol*. 1996;25(2):435–442.
4. Thygesen LC, Ersboll AK. When the entire population is the sample: strengths and limitations in register-based epidemiology. *Eur J Epidemiol*. 2014;29(8):551–558.
5. Liao JJ. Sample size calculation for an agreement study. *Pharm Stat*. 2010;9(2):125–132.
6. Krebs L, Langhoff-Roos J. Validation of registries: a neglected, but indispensable investment. *Paediatr Perinat Epidemiol*. 2014;28(5):351–352.
7. Lynge E, Sandegaard JL, Rebolj M. The Danish National Patient Register. *Scand J Public Health*. 2011;39(7 Suppl):30–33.
8. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159–174.
9. Lain SJ, Hadfield RM, Raynes-Greenow CH, et al. Quality of data in perinatal population health databases: a systematic review. *Med Care*. 2012;50(4):e7–e20.
10. Dansk Selskab for Obstetrik og Gynækologi. Anbefalede obstetriske koder [Recommended coding for obstetrics] [report on the Internet]; 2013. Available from: http://static.squarespace.com/static/5467abcce4b056d72594db79/546e7741e4b0d969a4f6c8e0/546e7745e4b0d969a4f6ca68/1385584564000/Obstetriske-diagnoser.pdf?format=original. Accessed January 26, 2015. Danish.

Paper V

BMJ Open

Effect of antenatal education in small classes versus standard auditorium-based lectures on use of pain relief during labour and of obstetric interventions: results from the randomised NEWBORN trial

Carina Sjöberg Brixval,[1] Lau Caspar Thygesen,[1] Solveig Forberg Axelsen,[1] Christian Gluud,[2] Per Winkel,[2] Jane Lindschou,[2] Tom Weber,[3] Pernille Due,[1] Vibeke Koushede[1]

To cite: Brixval CS, Thygesen LC, Axelsen SF, *et al*. Effect of antenatal education in small classes versus standard auditorium-based lectures on use of pain relief during labour and of obstetric interventions: results from the randomised NEWBORN trial. *BMJ Open* 2016;**6**:e010761. doi:10.1136/bmjopen-2015-010761

► Prepublication history and additional material is available. To view please visit the journal (http://dx.doi.org/10.1136/bmjopen-2015-010761).

Received 6 December 2015
Revised 18 May 2016
Accepted 19 May 2016

For numbered affiliations see end of article.

Correspondence to
Carina Sjöberg Brixval;
cabr@niph.dk

ABSTRACT

Objective: To examine the effect of an antenatal education programme in small classes versus standard auditorium-based lectures.

Design: Randomised trial using random-generated web-based 1:1 allocation.

Setting: The largest birth site in the Capital Region of Denmark, from August 2012 to May 2014.

Participants: 1766 pregnant women. Inclusion criteria ≥18 years, pregnant with a single child, and able to speak and understand Danish. Women were enrolled in the trial from 10+0 to 20+0 weeks of gestation.

Interventions: The intervention programme consisted of three times 2.5 hours of antenatal education in small classes (n=6–8 women), and focused on improving information and problem-solving skills for expectant parents in order to ease birth and the transition to parenthood. The control group received standard auditorium-based lectures consisting of two times 2 hours in an auditorium with participation of ~250 people.

Main outcome measures: The primary trial outcome was use of epidural analgesia. Other types of pain relief and obstetric interventions were analysed as explorative outcomes.

Results: There was no statistically significant difference in use of epidural analgesia between participants in the intervention group (30.9%) versus the control group (29.1%), adjusted OR 1.10 (95% CI 0.87 to 1.34). Also, the two groups did not differ regarding other types of pain relief or obstetric interventions. Concomitant birth preparation was common in both groups and highest in the control group, but did not seem to influence our results noticeably.

Conclusions: Antenatal education in small groups versus standard auditorium-based lectures did not differ regarding use of epidural analgesia, other pain relief, or obstetric interventions.

Trial registration number: NCT01672437; Results.

Strengths and limitations of this study

- This is the largest randomised trial evaluating the effect of a structured antenatal education programme in small classes.
- We developed a programme which could be implemented in the clinical setting if proven effective and compared the programme with standard care at the largest birth site in Denmark.
- We used proper methods for reducing the risks of bias; adequate sequence generation; allocation concealment; and use of an objectively measured primary outcome, epidural analgesia, reducing the risk of bias due to non-blinding.
- Attrition was low and evenly distributed between the groups.
- A total of only 19.6% of the invited women were accepted and randomised. These women differed from the general population regarding educational level and parity. This limits the generalisability of the trial results.

INTRODUCTION

Antenatal education has the aim to provide expectant parents with strategies for dealing with pregnancy, childbirth and parenthood.[1] Offers of antenatal education have undergone marked changes over time without evidence of the effect of various types of antenatal education on relevant outcomes, for example, outcomes related to birth.[2] A recent systematic review concluded that insufficient evidence exist as to whether antenatal education in small classes has an effect on obstetric or psychosocial outcomes.[3]

Epidural analgesia provides effective pain relief but is associated with adverse effects on the birth process, for example, longer second stage of labour,[4 5] and increased risk of operative birth[4–10] and of caesarean section.[5 10 11] Use of epidural analgesia and obstetric interventions, for example, vacuum extraction and caesarean section also have economic impacts on the health system.[12]

Women who are anxious during labour may be at increased risk of use of epidural analgesia as pain relief due to several mechanisms; they often seek admission to the labour ward at an earlier phase of labour;[13] anxiety and fear increase the risk of a longer active labour phase due to inhibited uterine contractility;[14] and anxious women perceive labour as more painful[6] and therefore receive more pain relief.[15] Also, childbirth self-efficacy may play an important role in the risk of receiving pain relief. Women with higher levels of self-efficacy and confidence in their ability to cope with birth report lower levels of anxiety,[16] perceive pain as less intensive,[17] and use less epidural analgesia during labour.[18] Childbirth self-efficacy may influence the timing of arrival to the labour ward, and it has been suggested that women with increased ability to cope with the early phase of labour will tend to arrive later at the labour ward.[19 20]

The promotion of self-efficacy beliefs during pregnancy may reduce anxiety, and this could possibly be provided through antenatal education in small classes. Antenatal education in small groups may provide an environment with the possibility for women to hear other pregnant women's experiences and for being encouraged by the other participants and the group facilitator. By these means, antenatal education in small classes may increase the woman's trust in her ability to cope with early labour,[21] and thereby reduce the likelihood of early admission[19] and decrease anxiety[21] and experience of labour pain.[22] This may in turn reduce the use of pain relief and obstetric interventions.

Antenatal education in small classes may, in addition, increase the women's knowledge uptake due to the possibility of being actively involved in the learning process.[23]

Few randomised trials have examined the effect of attending antenatal education in small groups compared with other forms of education on outcomes such as the use of pain relief or obstetric interventions,[19 24 25] and among these trials conclusions are conflicting.[3]

Owing to the sparse evidence from randomised trials, research about the effects of antenatal education in small classes on birth-related outcomes is still needed.[3] We, therefore, conducted a randomised trial to examine the effect of a general antenatal education programme in small classes versus standard education carried out as auditorium-based lectures. In this paper, we report the effect of the intervention on the primary outcome of the trial: use of epidural analgesia as well as the explorative outcomes: other types of pain relief and obstetric interventions.

METHODS

The NEWBORN trial is an individually randomised trial. The trial is registered at ClinicalTrials.gov (ID: NCT01672437), and a detailed description is published in our design article[26] and in our trial protocol.[27]

The NEWBORN trial took place at the largest birth site in Denmark, Hvidovre Hospital, situated in the Copenhagen Capital Region. More than 6500 deliveries take place at Hvidovre Hospital each year and the catchment area comprises a diverse population regarding sociodemographic characteristics.

Participants

Women were enrolled in the trial from 10+0 to 20+0 weeks of gestation. Inclusion criteria were expectant women, ≥18 years old at enrolment, singleton pregnancy, due to give birth at Hvidovre Hospital, and having the ability to speak and understand Danish. Exclusion criterion was not providing signed informed consent. Pregnant women were recruited from August 2012 to May 2014. The women received a written invitation to participate in the trial prior to their first visit to the midwife along with an informed consent form. Invitations were followed up by a phone call from a project employee. Initially, only primiparous women were eligible for participation, but due to slow recruitment also multiparous women were included ~6 months into the recruitment period in order to ensure adequate statistical power.[27] This change in the inclusion criteria was reported to ClinicalTrials.gov.

Randomisation

Baseline data were collected using a web-based questionnaire prior to randomisation. A project employee performed individual web-based randomisation at The Copenhagen Trial Unit according to a computer-generated allocation sequence of 1:1 with varying block sizes concealed to the investigators. The allocation was stratified for parity (primiparous or multiparous) and vulnerability (yes or no as defined by their general practitioner at the first pregnancy consultation in gestation week 6–10). There were eight criteria listed for vulnerability, for example; former or current psychiatric disorder, adverse psychosocial background, or concerns about parenting skills. The general practitioner categorised the women as vulnerable if she met one or more of these criteria. For non-vulnerable women, the block sizes used for randomisation were 10 and 20, for vulnerable women the block sizes were 4 and 6. These block sizes were used for primiparous as well as multiparous women. All the citizens in Denmark have a unique personal identification (CPR) number and the randomisation programme was set up to confirm the existence of the CPR number.

Intervention group

Women in the intervention group received an antenatal education programme—the NEWBORN programme—

focused on parental resources important for the birth process and for parenting. In short, the programme aimed at strengthening relationships and improving information and problem-solving skills for expectant parents in order to ease birth and the transition to parenthood. The woman's partner was also invited to participate in the programme. The consulting midwife and the facilitating midwife encouraged participation of the partner. Throughout the programme, there was a focus on increasing self-efficacy in relation to the different topics touched upon, for example, childbirth self-efficacy. Also, the programme aimed at enhancing social network among the participants and highlighted the importance of partner support. The programme was designed based on the recommendations for antenatal care from the Danish Health Authority[28] and developed using the Intervention Mapping approach.[29] A working group consisting of midwifes, health visitors, psychologists and family therapists delivered inputs for the form and content of the programme.

A detailed description of the programme has been presented elsewhere[26] and can be found as online supplementary material. Briefly, 110 groups of ~6–8 pregnant women and their partners met three times during pregnancy (gestation weeks 25, 33 and 35) for the duration of 2.5 hours per session. The sessions included among other things information and discussions about emotions and expectations related to the birth process, including information on pain relief and obstetric interventions. Accordingly, each woman was exposed to small group education for 7.5 hours during pregnancy.

The session in the 33rd week of gestation focused on pain relief and the birth process. The aim of this session was to provide the participants with information, and enhancing their existing knowledge and understanding of the normal course of labour, pain relief, and what might be expected if an obstetric intervention is necessary. For the topic on pain relief, the women were asked to discuss their thoughts and previous experiences with coping with pain and physical and mental strain, and to consider whether they could use any of these strategies during labour. Next, they discussed their thoughts and knowledge on various methods of pain relief. Plenary discussions and summing up thoughts and ideas were used so that participants could learn from and be inspired by one another. These methods were used to enhance the women's childbirth self-efficacy.

Also, the participants had access to a patient network website specifically developed for the NEWBORN trial to gain further information, communicate with other participants in the trial, and consult online with a midwife and a health visitor. At each session, the participants were encouraged by the instructors to use the website.

A total of 25 midwives with varying professional seniority and teaching experience facilitated the 110 classes. They enrolled for teaching themselves and were not specifically selected by the trial investigators. The instructors followed a detailed teaching manual developed for the trial.[30]

In the Copenhagen area, different kinds of birth and parent education offers are provided by private stakeholders. These offers include, for example, mindfulness training, physical exercise training and mental preparation for delivery. Participants in the intervention group were permitted to make use of concomitant birth and parent education.

Control group

Women in the control group received the standard education offered from Hvidovre Hospital consisting of two antenatal lectures of 2 hours, each on birth and breast feeding in an auditorium with participation of up to 250 people. Accordingly, each woman was exposed to large group education for 4 hours. The content of the lecture on birth included information on, for example, what to do at home when labour had begun; information on the location of labour ward at the hospital; phases of the labour and information on different types of pain relief. The form was passive information giving from a midwife to the participants in the lectures.

Midwives conducting the lectures volunteered for the teaching. To avoid contamination of conditions, midwives facilitating the group-based experimental programme were not allowed to teach the antenatal lectures in the control group.

In addition to participants in the intervention group, participants in the control group were permitted to make use of concomitant birth and parent education.

Blinding

It was not possible to blind participants or service providers. The outcome assessors; midwives, and physicians at the labour ward were not informed about the women's participation in the trial. Data were blinded by a data manager and the investigators were therefore blinded to participants' intervention category during data assessment and analyses. Participants' intervention category was not revealed to the investigators until the Steering Committee of the trial had drawn two conclusions about intervention effects on outcomes under code.[31 32]

Variables

The primary outcome of the trial was the use of epidural analgesia during labour. The use of other types of pain relief and obstetric interventions was examined as explorative outcomes.

Data on the use of pain relief, obstetric interventions, and other variables related to the birth was assessed from the hospital-based register at Hvidovre Hospital, the Obstetric Database. All births performed at Hvidovre Hospital and two other birth sites in the Capital Region are included in this database and entries are made by CPR number. No information about birth was collected specifically for the NEWBORN trial. In a validation

study, we found that the validity of information on epidural analgesia and selected obstetric interventions was high in the obstetric database when using the medical records as the gold standard.[33] The positive predictive values for epidural analgesia, vacuum extraction and emergency caesarean section were 1.00, and for elective caesarean section the positive predictive value was 0.96.[33]

The following variables were used for examination of baseline differences: *Educational level* was measured by the question: 'What is your highest completed education'? The educational level was dichotomised into ≤medium tertiary education versus higher tertiary education. *Body mass index* (kg/m^2) was calculated using the information on prepregnancy weight and height reported by the woman at the first pregnancy consultation at the general practitioner. *Living with child's father* was self-reported by ticking the response category 'Living with the child's father' in the question: 'Which grown-ups do you live with'? *Planned pregnancy* was self-reported by the question: 'Is this pregnancy planned, partly planned or not planned' and dichotomised into: planned (yes or partly) versus not planned. *Self-rated physical and mental health* was measured by the items: 'How would you describe your physical/mental health status altogether'? Response categories: 'Excellent, very good, good, poor, very poor'. Self-rated physical/mental health was dichotomised into excellent/very good versus good, poor, very poor. *Feeling of stress* was measured by the item: 'Do you feel stressed'? Response categories: 'no; yes, a little; yes, moderately; yes, a lot'. Stress was dichotomised into no versus yes, a little; yes, moderately; yes, a lot. Antenatal depressive symptomatology was measured by the Edinburgh Postnatal Depression Scale[34] posed in the baseline questionnaire. A score of 13 or more were categorised as antenatal depressive symptomatology. Perceived stress was measured by the Perceived Stress Scale.[35]

In order to give an indication of the quality of the delivery of the programme, we assessed adherence to the programme by tablet-based questionnaires. After each session, the participants were asked whether they had been through the topics of the day. For example, after session 2, the participants were asked: 'Have you heard about "coping with pain and pain relief" today'. Participants could answer 'yes', 'no', or 'don't know'.

Data on use of concomitant birth and parent education were collected by questionnaires at gestation week 37 and 9 weeks after birth. We examined the prevalence of antenatal depressive symptomatology among participants as a potential adverse outcome. Participants in the intervention group could potentially have experienced more antenatal depressive symptoms, for example, due to a raised awareness on couple communication and potential relationship problems through the sessions. Antenatal depressive symptomatology was measured by the Edinburg Postnatal Depression Scale[34] collected by questionnaire in gestation week 37. Although initially developed for measuring depressive symptoms in the postnatal period, the scale has been validated for use during pregnancy as well.[36] Women with a score of 13 or more were categorised with antenatal depressive symptomatology as recommended in a former Swedish study.[36]

Sample size

The sample size calculation was based on the primary outcome of the trial, use of epidural analgesia. Previous data from trials[19] and hospital registers[37] indicate that the proportion of women who use epidural analgesia is between 23% and 41%. We assumed that 31% in the control group would receive epidural analgesia and that this proportion could be reduced to 25% in the intervention group (a relative risk reduction of 19%). Our original sample size calculation was based on a power of 0.90 and a significance level of 0.05 requiring randomisation of 2350 women to detect significant intervention effects. However, due to slow recruitment power was reduced to 0.80 requiring randomisation of 1756 women. This sample size adjustment was carried out after inclusion of 1050 participants without inspection of the data.[27 38]

Statistical analysis

Data were planned to be analysed according to the intention-to-treat principle and following the recommendations of the CONSORT statement.[39 40]

Main analyses

Differences in frequency of use of epidural analgesia, other types of pain relief, and obstetric interventions between the two groups were tested in logistic regression models adjusted for the protocol-specified stratification variables; parity and vulnerability. ORs and 95% CIs, as well as relative risk (RR) estimates with 95% CI were calculated. Difference in the proportion of the adverse outcome antenatal depressive symptomatology between the groups was tested by χ^2 test.

Handling of missing data

We tested whether missing values of the primary outcome, epidural analgesia, were missing completely at random (MCAR) by Little's test.[41] Also, 'worst case' and a 'best case' scenario analyses of the potential impact of missing values were conducted. In the worst-case scenario, missing values of epidural analgesia in the intervention group were imputed by a 'yes' and missing values of the control group were imputed by a 'no'. In the best-case scenario, missing values of epidural analgesia in the intervention group were imputed by a 'no' and missing values of the control group were imputed by a 'yes'.

We selected participants with full report on the primary outcome for the modified intention-to-treat analysis (see results).

Sensitivity and per-protocol analyses

We conducted a post hoc analysis with the aim of examining the impact of concomitant birth and parent preparation on the primary outcome. From the modified intention-to-treat cohort, we excluded the participants who made use of concomitant birth and parent education in both intervention groups.

The compliance with the randomised interventions was not 100%. We therefore planned per-protocol analyses in our trial protocol. Definition of per-protocol conditions were carried out prior to data analysis. The results from the per-protocol analyses are interpreted as explorative. We compared the use of epidural analgesia between the two intervention groups in per-protocol populations defined as follows:

1. Participants in the intervention group who participated in all three sessions before birth and used the website at least once versus all participants in the control group were selected from the modified intention-to-treat cohort.
2. Participants in the intervention group who participated in all three sessions before birth and used the website at least once versus participants in the control group who participated in both antenatal lectures were selected from the modified intention-to-treat cohort.

All statistical analyses were performed using SAS V. 9.3, SAS Institute Inc. The level of significance was set to 0.05.

Table 1 Baseline characteristics of women enrolled in the NEWBORN trial (n=1766)

	Experimental (n=883)	Control (n=883)
Age at birth in years (mean (SD))*	30.7 (4.1)	30.8 (4.1)
Nulliparous, n (%)	787 (89.1)	785 (88.9)
Vulnerable women, n (%)	42 (4.8)	42 (4.8)
Educational level (higher tertiary education), n (%)	659 (75.6)	663 (76.5)
Body mass index kg/m^2 (mean (SD))*	23.4 (4.0)	23.3 (4.1)
Living with child's father (yes), n (%)	828 (93.8)	848 (96.0)
Planned pregnancy (yes/partly), n (%)	801 (90.9)	808 (91.5)
Self-rated physical health status (excellent/very good), n (%)	605 (68.6)	628 (71.2)
Self-rated mental health status (excellent/very good), n (%)	635 (72.0)	669 (75.9)
Not feeling stressed, n (%)	425 (48.2)	433 (49.2)
Edinburgh Postnatal Depression Scale score of 13 or more, n (%)	42 (4.8)	28 (3.2)
Perceived Stress Scale score (mean (SD))	12.5 (5.2)	12.2 (5.2)

*Based on women with birth data (n=1711).

RESULTS

Participant flow and baseline data

During the recruitment period, 8997 women were invited to participate in the NEWBORN trial. Of these, 1766 women (19.6%) accepted participation and were randomised—883 women to the intervention group versus 883 to the control group. At baseline, the characteristics among the intervention and control groups seem well balanced (table 1).

The attrition was similar in the two groups (figure 1). Little's test for MCAR was insignificant (p=0.64). Therefore, no imputation of missing values was performed. The modified intention-to-treat analysis therefore included 1711 participants (858 in the intervention group vs 853 in the control group).

Effect of the experimental intervention

We found no effect of the NEWBORN intervention. Among women in the intervention group, 30.5% received epidural analgesia compared with 29.1% in the control group (adjusted OR=1.10 (0.87 to 1.34), p=0.41). None of the exploratory outcomes differed statistically between the two groups (table 2). We found no adverse effects of attending the experimental group on antenatal depressive symptomatology. The proportion of participants categorised as having antenatal depressive symptomatology at gestation week 37 was 5.6% in the intervention group and 6.8% in the control group (p=0.34).

We conducted 'worst-case' and 'best-case' scenario analyses to assess the potential impact of missing values. Results from the best-case scenario showed no difference between intervention group and control group on the use of epidural analgesia (adjusted OR=0.93 (0.76 to 1.14), p=0.49). In the worst-case scenario, the results indicated a negative impact of the intervention (adjusted OR=1.25 (1.02 to 1.54), p=0.03) (see online supplementary table S1).

Adherence to the programme in session 2

To give an indication of the quality of delivery of the programme, we assessed the facilitator's adherence to the programme content in session 2. Adherence was reported high by the participants. More than 97% of the participants reported to have heard about the topics: 'expectations in relation to birth', 'what to do at home in the early phase of labour', 'the normal course of labour, pain relief and coping strategies', and 'partner support during labour'. A total of 88% of the participants reported having been through the topic 'when there is a need to intervene in labour'.

Use of birth and parent education offers

Use of birth and parent education offers was unequally distributed among intervention groups (see online supplementary table S2). There were a considerably higher

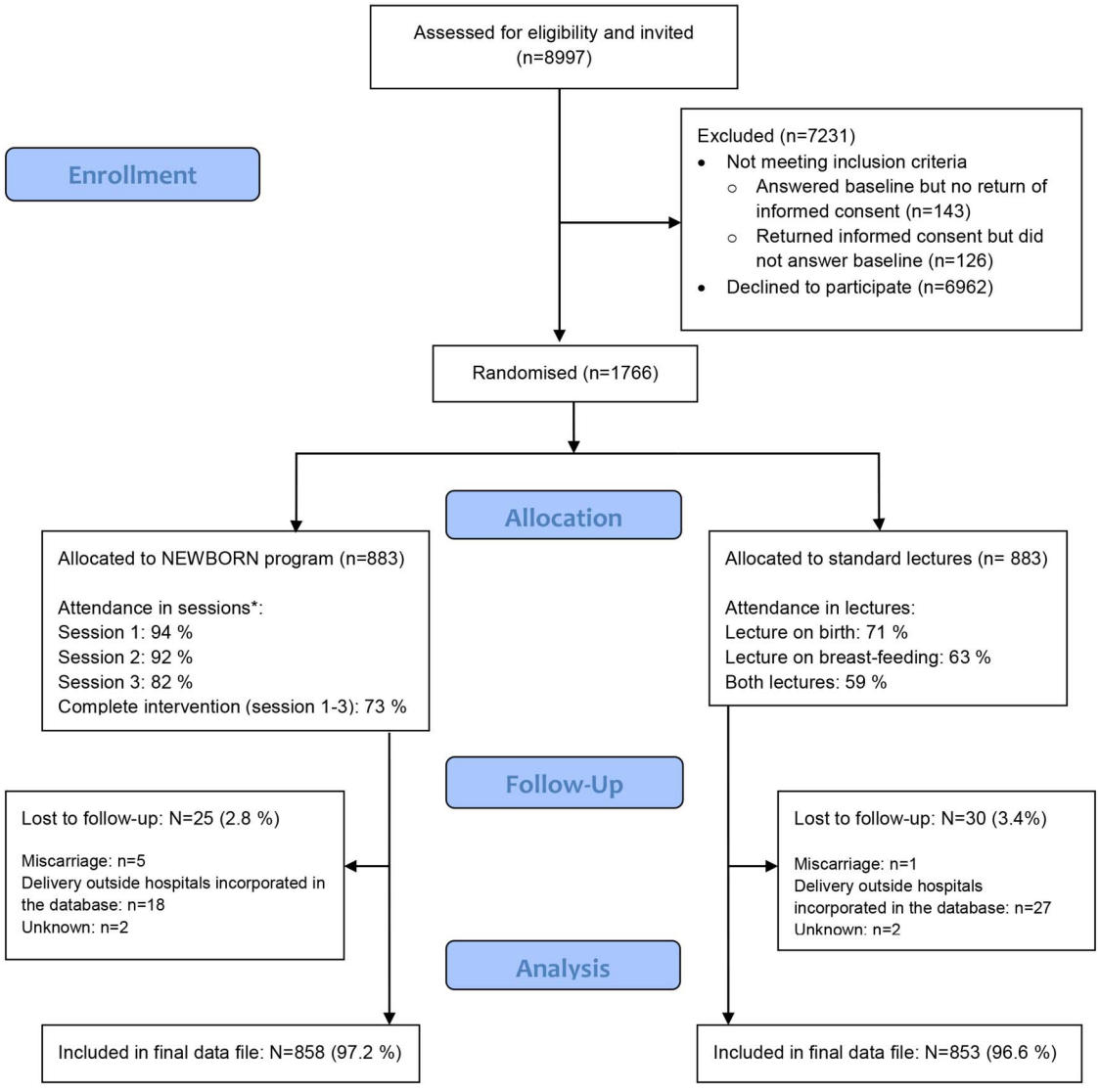


Figure 1 Flow diagram of recruitment, randomisation and participation in the NEWBORN trial.

proportion of participants in the control group (38.7%) who attended other types of birth and parent preparation offers than among participants in the intervention group (25.0%). Also, there were more participants who did not attend any birth and parent education offers in the control group (11%) than among participants in the intervention group (2.5%).

Additional analyses

We performed an additional sensitivity analysis examining the effect of the intervention on the use of epidural analgesia excluding women who made use of concomitant preparation education. This reduced the sample from 1711 women to 1052 women. Results were similar to the results from the modified intention-to-treat analysis, that is, there was no effect of the intervention (table 3). In the per-protocol analyses where we examined the effect of the intervention among participants adhering to the intervention, the sample was reduced with ~25%. Also, results from these analyses were consistent with the results from the modified intention-to-treat analysis (table 3).

DISCUSSION

The results from this randomised trial showed that the experimental education consisting of small classes for 7.5 hours versus control education with large group lectures for 4 hours gave no difference in the use of epidural analgesia, other types of pain relief during labour or obstetric interventions. Use of private birth and parent preparation offers were considerably higher

Table 2 ORs (95% CI) and relative risks (RR) (95% CI) for use of pain relief and obstetric interventions when comparing the experimental programme with standard lectures

	Experimental	Control	Adjusted OR (95% CI)*	p Value*	Crude OR (95% CI)	p Value	Crude RR (95% CI)	p Value
Pain relief								
Epidural analgesia	265/858 30.9%	248/853 29.1%	1.10 (0.87 to 1.34)	0.41	1.09 (0.89 to 1.34)	0.41	1.07 (0.92 to 1.24)	0.37
Pudendal nerve block	79/858 9.2%	64/853 7.5%	1.25 (0.89 to 1.77)	0.20	1.25 (0.89 to 1.76)	0.20	1.23 (0.90 to 1.68)	0.19
Water immersion	157/858 18.3%	148/853 17.4%	1.07 (0.83 to 1.37)	0.61	1.07 (0.83 to 1.37)	0.60	1.06 (0.86 to 1.30)	0.57
Acupuncture	115/858 13.4%	116/853 13.6%	0.98 (0.74 to 1.30)	0.90	0.98 (0.75 to 1.30)	0.91	0.99 (0.78 to 1.26)	0.94
Intracutaneous sterile water injection	74/858 8.6%	80/853 9.4%	0.91 (0.65 to 1.27)	0.58	0.91 (0.66 to 1.27)	0.59	0.93 (0.68 to 1.25)	0.61
Morphine	62/858 7.2%	48/853 5.6%	1.31 (0.89 to 1.94)	0.18	1.31 (0.89 to 1.93)	0.18	1.29 (0.90 to 1.86)	0.17
Nitrous oxide	4/858 0.5%	8/853 0.9%	0.50 (0.15 to 1.66)	0.25	0.50 (0.15 to 1.65)	0.25	0.50 (0.15 to 1.65)	0.26
Obstetric interventions								
Vacuum extraction	132/858 15.4%	127/853 14.9%	1.04 (0.80 to 1.36)	0.78	1.03 (0.80 to 1.35)	0.78	1.04 (0.83 to 1.30)	0.74
Emergency caesarean section	149/858 17.4%	147/853 17.2%	1.01 (0.78 to 1.30)	0.94	1.01 (0.79 to 1.30)	0.94	1.01 (0.82 to 1.25)	0.90
Elective caesarean section	34/858 4.0%	42/853 4.9%	0.80 (0.50 to 1.27)	0.34	0.80 (0.50 to 1.27)	0.34	0.81 (0.52 to 1.26)	0.35

Analyses are based on the modified intention-to-treat population (N=1711).
*Adjusted for trial stratification variables: vulnerability and parity.

Table 3 ORs (95% CI) and relative risks (RRs) (95% CI) for use of epidural analgesia when comparing the experimental programme to standard lectures in additional analyses

	Experimental	Control	Adjusted OR (95% CI)* p Value	Crude OR (95% CI) p Value	Crude RR (95% CI) p Value
Population					
Exclusion of participants who made use of concomitant birth and parent education, N=1052 (573 in intervention group vs 479 in control group)	31.8% (182/573)	29.4% (141/479)	1.08 (0.83 to 1.41) 0.59	1.11 (0.87 to 1.41) 0.42	1.09 (0.91 to 1.26) 0.42
Per-protocol analysis #1, N=1251 (398 in intervention group vs 853 in control group)	30.7% (122/398)	29.1% (248/853)	1.02 (0.79 to 1.33) 0.88	1.08 (0.83 to 1.40) 0.57	1.05 (0.88 to 1.26) 0.57
Per-protocol analysis #2, N=1205 (398 in intervention group vs 532 in control group)	30.7% (122/398)	32.5% (173/532)	0.93 (0.70 to 1.23) 0.59	0.92 (0.69 to 1.21) 0.55	0.94 (0.78 to 1.14) 0.55

#1: Participants in the intervention group who participated in all three sessions before birth and used the website at least once versus all participants in the control group.
#2: Participants in the intervention group who participated in all three sessions before birth and used the website at least once versus participants in the control group who participated in both antenatal lectures.
*Adjusted for trial stratification variables: vulnerability and parity.

among participants in the control group compared with participants in the intervention group, but also no participation in birth and parent preparation was more frequent in the control group than among participants in the intervention group. We examined the impact of the concomitant education by excluding women that participated in other education and found that this exclusion did not alter our results noticeably.

We hypothesised that the NEWBORN programme would increase childbirth self-efficacy and by this reduce the use of pain relief. We have examined the effect of the programme on the intermediate trial outcome; childbirth self-efficacy measured by three single items developed for the NEWBORN trial. In the intervention group, 4.1% of the women had low confidence in their own ability to cope with early phase of labour before going to the labour ward compared with 8.0% in the control group. Fewer women in the intervention group (5.0%) felt low confidence in their own ability to handle the birth process no matter how it turns out compared with the control group (7.4%).[42] Hence, these results suggest that although the intervention had no effect on the epidural analgesia, the programme may have the potential to enhance the women's childbirth self-efficacy. Former studies have found that women with higher levels of self-efficacy perceive pain as less intensive[17] and use less epidural analgesia during labour[18] compared with women with lower levels of self-efficacy. The potential associations between childbirth self-efficacy and experience of pain and use of epidural analgesia have not been investigated in the present study.

Only three randomised trials have examined the effect of attending antenatal education in small groups compared with other forms of education on outcomes, such as the use of pain relief or obstetric interventions.[19 24 25] Two of these trials were performed among women screened positive for fear of childbirth limiting generalisation of results to the general population.[24 25] One former Danish trial[19] examined the effect of antenatal education classes versus no education among 1193 primiparous women recruited among a diverse population group not limited to a high-risk population. This trial by Maimburg *et al*[19] is comparable to our NEWBORN trial regarding the included population, but they compared small classes versus no intervention. Maimburg *et al*[19] reported a statistically significant reduced use of epidural analgesia in their experimental group, but not of other types of pain relief and obstetric interventions. The two trials differ regarding the control intervention and we included primiparous and multiparous women, whereas Maimburg *et al* only included primiparous. Furthermore, we used 25 voluntary midwives with varying teaching experience, whereas in the trial by Maimburg *et al* classes were taught by four selected midwives. Also, the midwives in the Maimburg trial may have gained more teaching experience during the trial period compared with the midwives in the NEWBORN trial, as some of the midwives in our trial taught only a few classes.

Strengths and limitations

This randomised trial is to our knowledge the largest trial assessing the effect of antenatal education in small classes versus auditorium-based lectures. The intervention was developed using a systematic framework for health promotion programme planners.[29] This systematic framework aids effective decision-making at each step in intervention planning, implementation and evaluation. We focused on conducting a trial using standard care as control condition and tested a birth and parent preparation programme that would be feasible to implement in an everyday clinical practice setting if proven effective. We chose a control condition that is relevant to public health; standard care instead of a different antenatal intervention, and the study population was recruited among a diverse population group and not limited to a high-risk population. The attrition was low (3%) and distributed evenly between the intervention and control groups.

It was not possible to blind participants or educators which may introduce bias. However, using an objective primary outcome, such as epidural analgesia, reduces the risk of bias due to lack of blinding.[43 44] The outcome assessors, midwives at the labour ward, were not informed about the women's participation in the trial but it cannot be ruled out that the women informed the midwife about their intervention status. However, we consider it unlikely that the decision to provide pain relief or perform obstetric interventions rely on the intervention status as such decisions are made by the midwives and physicians at the labour ward.

Initially, only primiparous women were eligible for inclusion in the trial. During the recruitment period, we allowed for inclusion of multiparous women. This was carried out for practical reasons to ensure adequate power in the analyses. Although this change was reported to the clinical trial register, the posterior inclusion of multiparous women must be considered a limitation.

Of the 8997 pregnant women invited to participate in the trial, only 19.6% were accepted and were randomised. Although we aimed to recruit a diverse population group to the NEWBORN trial, the participants were predominantly primiparous women and women with a higher education level compared to the general population of Copenhagen women in the same age group.[45] The high proportion of women with a university education in the trial population may imply that the women included in the trial find this teaching form more appealing than the general population. Moreover, the proportion receiving pain relief and obstetric interventions (except elective caesarean section) were higher among the trial population than among the total population of women giving birth at Hvidovre Hospital.[46] These discrepancies between the trial population and background population characteristics may limit the generalisability of the trial results, and the intervention might have different effects among multiparous or women of a lower educational level. It may be beneficial to conduct research focusing on the effect of the programme among subgroups, for example, women with lower educational level or vulnerable women. Also, further analyses taking adherence to the programme into consideration would contribute with more thorough knowledge of the impact of the programme. These issues need to be investigated before recommendation of implementation of the programme in clinical practice can be validly expressed.

CONCLUSIONS

The results from the NEWBORN trial showed no difference in use of epidural analgesia, other types of pain relief during labour, or obstetric interventions between women randomised to antenatal education in small classes versus standard lectures. The effects of the intervention on the secondary outcomes of the NEWBORN trial: perceived stress, parenting stress, and parenting alliance will be reported in later articles.

Author affiliations
[1]National Institute of Public Health, University of Southern Denmark, Copenhagen, Denmark
[2]Copenhagen Trial Unit, Centre for Clinical Intervention Research, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark
[3]Department of Obstetrics and Gynaecology, Hvidovre Hospital, Copenhagen University Hospital, Hvidovre, Denmark

Acknowledgements The authors would like to acknowledge all participants in the NEWBORN trial as well as the NEWBORN Trial Group[26] involved in the trial. They thank the Danish Cancer Society for funding of the trial. Moreover, they also thank Janus Engstrøm, Copenhagen Trial Unit, for development of the randomisation system and provision of allocation sequence; Steen Christian Rasmussen, Hvidovre Hospital, for extraction of data from the Obstetric Database and Bjarne Laursen, National Institute of Public Health, for valuable help with data management.

Contributors CSB, SFA, PD and VK developed the study design in collaboration with TW, JL and CG. PW wrote the statistical analysis plan. CSB, SFA and VK collected the data. CSB and LCT conducted the statistical analyses. CSB drafted the manuscript. All authors have been involved with interpretation of results, critical revision of the manuscript and final approval of the submitted manuscript.

Funding This work was supported by the Danish Cancer Society. The funders have not been involved in trial design, data collection and analysis, or decision to publish this manuscript.

Competing interests None declared.

Ethics approval The NEWBORN trial has been assessed by and registered in the Capital Region's ethics committee 23 March 2012 (CVR/SE number 30113713) and registered and listed in the Danish Data Protection Agency (reference number 2011-54-1289).

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement The authors commit to making the relevant anonymised patient-level data available on reasonable request. Please contact the corresponding author.

REFERENCES

1. Ahlden I, Ahlehagen S, Dahlgren LO, *et al*. Parents' expectations about participating in antenatal parenthood education classes. *J Perinat Educ* 2012;21:11–17.
2. Gagnon AJ, Sandall J. Individual or group antenatal education for childbirth or parenthood, or both. *Cochrane Database Syst Rev* 2007;(3):CD002869.
3. Brixval CS, Axelsen SF, Lauemøller SG, *et al*. The effect of antenatal education in small classes on obstetric and psycho-social outcomes-a systematic review. *Syst Rev* 2015;4:1–9.
4. Anim-Somuah M, Smyth RM, Jones L. Epidural versus non-epidural or no analgesia in labour. *Cochrane Database Syst Rev* 2011;(12):CD000331.
5. Zimmer EZ, Jakobi P, Itskovitz-Eldor J, *et al*. Adverse effects of epidural analgesia in labor. *Eur J Obstet Gynecol Reprod Biol* 2000;89:153–7.
6. Jones L, Othman M, Dowswell T, *et al*. Pain management for women in labour: an overview of systematic reviews. *Cochrane Database Syst Rev* 2012;3:CD009234.
7. Poignant M, Hjelmstedt A, Ekeus C. Indications for operative delivery between 1999–2010 and induction of labor and epidural analgesia on the risk of operative delivery—a population based Swedish register study. *Sex Reprod Healthc* 2012;3:129–34.
8. Nguyen US, Rothman KJ, Demissie S, *et al*. Epidural analgesia and risks of cesarean and operative vaginal deliveries in nulliparous and multiparous women. *Matern Child Health J* 2010;14:705–12.
9. Kjaergaard H, Olsen J, Ottesen B, *et al*. Obstetric risk indicators for labour dystocia in nulliparous women: a multi-centre cohort study. *BMC Pregnancy Childbirth* 2008;8:45.
10. Eriksen LM, Nohr EA, Kjaergaard H. Mode of delivery after epidural analgesia in a cohort of low-risk nulliparas. *Birth* 2011;38:317–26.
11. Bannister-Tyrrell M, Ford JB, Morris JM, *et al*. Epidural analgesia in labour and risk of caesarean delivery. *Paediatr Perinat Epidemiol* 2014;28:400–11.
12. Tracy SK, Tracy MB. Costing the cascade: estimating the cost of increased obstetric intervention in childbirth using population data. *BJOG* 2003;110:717–24.
13. Carlsson IM, Hallberg LR, Odberg Pettersson K. Swedish women's experiences of seeking care and being admitted during the latent phase of labour: a grounded theory study. *Midwifery* 2009;25:172–80.
14. Lederman RP, Lederman E, Work BA Jr, *et al*. The relationship of maternal anxiety, plasma catecholamines, and plasma cortisol to progress in labor. *Am J Obstet Gynecol* 1978;132:495–500.
15. Alehagen S, Wijma K, Wijma B. Fear during labor. *Acta Obstet Gynecol Scand* 2001;80:315–20.
16. Lowe NK. Self-efficacy for labor and childbirth fears in nulliparous pregnant women. *J Psychosom Obstet Gynaecol* 2000;21:219–24.
17. Lowe NK. The pain and discomfort of labor and birth. *J Obstet Gynecol Neonatal Nurs* 1996;25:82–92.
18. Carlsson IM, Ziegert K, Nissen E. The relationship between childbirth self-efficacy and aspects of well-being, birth interventions and birth outcomes. *Midwifery* 2015;31:1000–7.
19. Maimburg RD, Vaeth M, Durr J, *et al*. Randomised trial of structured antenatal training sessions to improve the birth process. *BJOG* 2010;117:921–8.
20. Beebe KR, Lee KA, Carrieri-Kohlman V, *et al*. The effects of childbirth self-efficacy and anxiety during pregnancy on prehospitalization labor. *J Obstet Gynecol Neonatal Nurs* 2007;36:410–18.
21. Taheri Z, Mazaheri MA, Khorsandi M, *et al*. Effect of educational intervention on self-efficacy for choosing delivery method among pregnant women in 2013. *Int J Prev Med* 2014;5:1247–54.
22. Ip WY, Tang CS, Goggins WB. An educational intervention to improve women's ability to cope with childbirth. *J Clin Nurs* 2009;18:2125–35.
23. Illeris K. *Contemporary theories of learning: learning theorists…in their own words*. Routledge, 2009.
24. Werner A, Uldbjerg N, Zachariae R, *et al*. Self-hypnosis for coping with labour pain: a randomised controlled trial. *BJOG* 2013;120:346–53.
25. Rouhe H, Salmela-Aro K, Toivanen R, *et al*. Obstetric outcome after intervention for severe fear of childbirth in nulliparous women—randomised trial. *BJOG* 2013;120:75–84.
26. Koushede V, Brixval CS, Axelsen SF, *et al*. Group-based antenatal birth and parent preparation for improving birth outcomes and parenting resources: study protocol for a randomised trial. *Sex Reprod Healthc* 2013;4:121–6.
27. Koushede V. Project Newborn–Preparation for Birth and Parenthood 2014. http://www.interventionsforskning.dk/userfiles/files/nyfdt/projektprotokol-version-23_2.pdf (accessed 7 Aug 2015).
28. Danish Health Authority. *Recommendations for Antenatal Care (in Danish)*. Copenhagen, 2009. http://sundhedsstyrelsen.dk/publ/Publ2009/CFF/gravide/Svangreomsorgen.pdf (accessed 9 Mar 2015).
29. Bartholomew LK, Parcel GS, Kok G, *et al*. *Planning health promotion programs: an intervention mapping approach*. John Wiley & Sons, 2011.
30. Axelsen SF, Brixval CS, Koushede V. Course material for teachers: NEWBORN—preparation for birth and parenthood. 2012.
31. Gotzsche PC. Blinding during data analysis and writing of manuscripts. *Control Clin Trials* 1996;17:285–90; discussion 90–3.
32. Jarvinen TL, Sihvonen R, Bhandari M, *et al*. Blinded interpretation of study results can feasibly and effectively diminish interpretation bias. *J Clin Epidemiol* 2014;67:769–72.
33. Brixval CS, Thygesen LC, Johansen NR, *et al*. Validity of a hospital-based obstetric register using medical records as reference. *Clin Epidemiol* 2015;7:509–15.
34. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. *Br J Psychiatry* 1987;150:782–6.
35. Cohen S, Kamarck T, Mermelstein R. A global measure of perceived stress. *J Health Soc Behav* 1983;24:385–96.
36. Rubertsson C, Borjesson K, Berglund A, *et al*. The Swedish validation of Edinburgh Postnatal Depression Scale (EPDS) during pregnancy. *Nord J Psychiatry* 2011;65:414–18.
37. Hvidovre Hospital. Births in numbers 2012 (in Danish). https://www.hvidovrehospital.dk/afdelinger-og-klinikker/foedeomraadet/om-afdelingen/Sider/Foedsler-i-tal.aspx (accessed 9 June 2013).
38. ClinicalTrials.gov. NewPreBP: Project Newborn—Preparation for Birth and Parenthood. https://clinicaltrials.gov/ct2/results?term=NCT01672437 Last update: December 2015 (accessed 3 Dec 2015).
39. Moher D, Hopewell S, Schulz KF, *et al*. CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. *J Clin Epidemiol* 2010;63:e1–37.
40. Boutron I, Moher D, Altman DG, *et al*. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med* 2008;148:295–309.
41. Little RJA. A test of missing completely at random for multivariate data with missing values. *J Am Stat Assoc* 1988;83:1198–202.
42. Brixval CS, Axelsen SF, Thygesen LC, *et al*. Antenatal education in small classes may increase childbirth self-efficacy: results from a Danish randomised trial. *Sex Reprod Healthc* 2016, in press. doi:10.1016/j.srhc.2016.03.003
43. Wood L, Egger M, Gluud LL, *et al*. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008;336:601–5.
44. Savovic J, Jones HE, Altman DG, *et al*. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. *Ann Intern Med* 2012;157:429–38.
45. Statistics Denmark. Population highest completed education 2014. http://www.statistikbanken.dk/statbank5a/default.asp?w=1422 (accessed 16 Mar 2015).
46. Hvidovre Hospital. Births in numbers 2014 (in Danish). https://www.hvidovrehospital.dk/afdelinger-og-klinikker/foedeomraadet/om-afdelingen/Sider/Foedsler-i-tal.aspx (accessed 7 Aug 2015).

BMJ Open

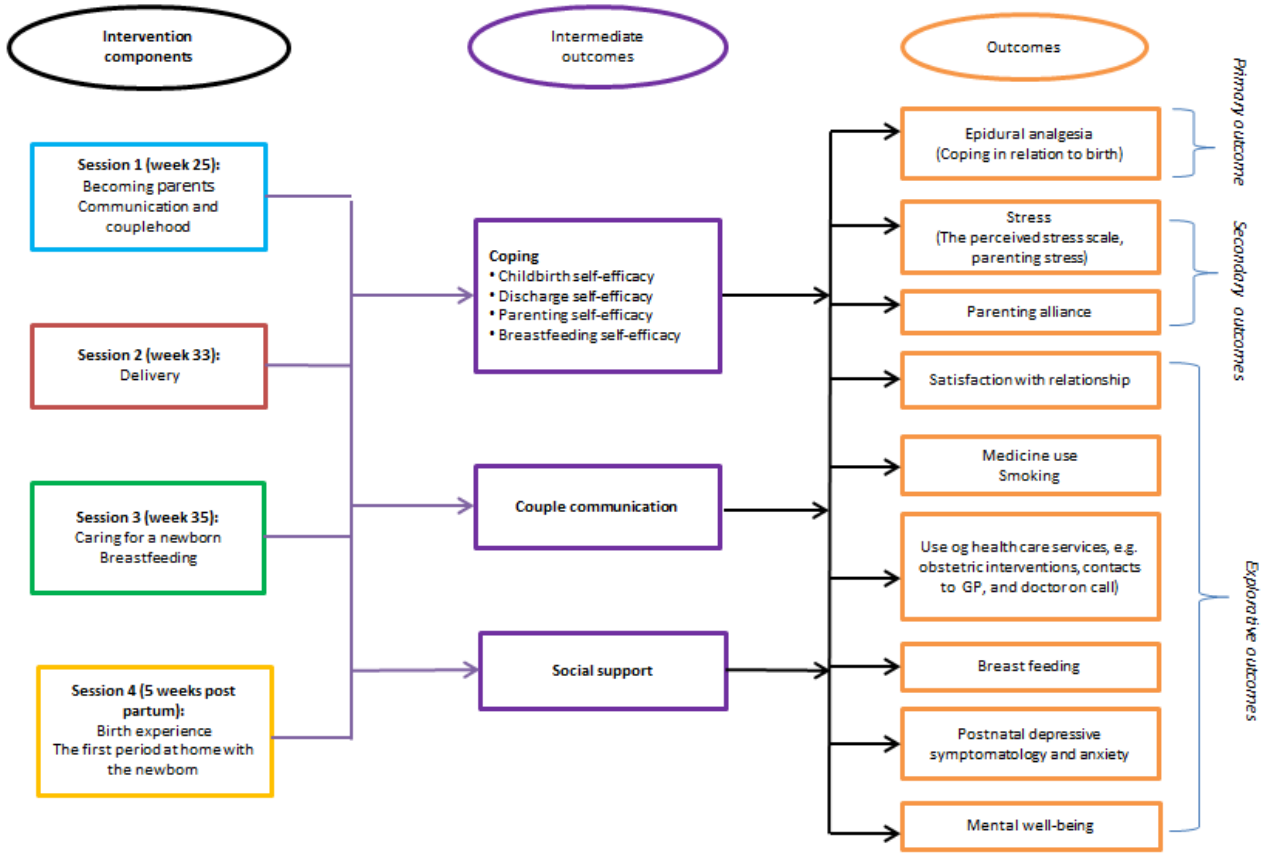
Effect of antenatal education in small classes versus standard auditorium-based lectures on use of pain relief during labour and of obstetric interventions: results from the randomised NEWBORN trial

Carina Sjöberg Brixval, Lau Caspar Thygesen, Solveig Forberg Axelsen, Christian Gluud, Per Winkel, Jane Lindschou, Tom Weber, Pernille Due and Vibeke Koushede

BMJ Open 2016 6:
doi: 10.1136/bmjopen-2015-010761

Supplementary material: Model of the program theory of the NEWBORN trial and the NEWBORN program in detail

Figure 1: Model of the program theory of the NEWBORN trial



Detailed description of the NEWBORN program

A detailed intervention manual was created for the trial facilitators. The following subjects were covered in the sessions:

The NEWBORN program included short verbal presentations from the group facilitator, individual exercises, short film presentations, time for discussions and reflection. Parents were given homework in the form of minor exercises in preparation to each session. Educational subjects were: the transition to parenthood;

couple communication; birth; breast feeding; and taking care of a newborn. A patient-network website was created as a supplement to the sessions. The program was focused on parenting resources important to the birth process, parenting, and mental health, and that appear amenable to change, i.e.: social support, parenting alliance and communication with partner, cognitive coping, e.g., self-efficacy and parenting skills. These elements have been addressed in the following manner:

I. Social support: formal and informal, emotional, informational and instrumental. Groups of 6-8 couples were offered three times 2.5 hour sessions during pregnancy and one session five weeks post-partum. The groups were composited to enable participants to establish relations with other expectant parents in their local area. Sessions were facilitated by a midwife and in the postnatal session a health visitor also participated. A patient-network website were provided with the purpose of enabling parents to gain further information, communicate with other parents and consult online with a midwife and a health visitor.

II. Parenting alliance: adding a component supporting the couples in the transition to parenthood and couple communication.

III. Cognitive coping: embedding sources of self-efficacy into program content and delivery, and by creating an environment which enables parents to discuss feelings and concerns, enhances their awareness of own resources, problem-solving strategies, and future challenges in parenting and emotional regulation.

IV. Parenting skills: increasing information and exercises with feedback, e.g., on recognizing signs and symptoms of thriving in the newborn, couple communication etc.

In short, the approach aimed at strengthening relationships and improving information and problem solving skills for expectant parents in order to ease birth and the transition to parenthood.

To maximize the potential for population uptake classes were established at three local midwifery sites. A comprehensive guide and education material for course facilitators were developed. The framework for the

classes was based on an estimate of adequate time allocated to each subject, and what service providers deemed a sustainable service.

Session 1 (approximately 25 weeks gestation):

• The transition to parenthood – new roles and responsibilities, emotional adjustment

• Common changes and challenges in the relationship during and after pregnancy

• Couple communication

• Meaning of own childhood when becoming parents

During the first session the parents introduced themselves to one another by a short exercise and were introduced to the scope and outline of the entire program. The midwife invited the parents to think about and since discuss their expectations of the greatest joys and greatest challenges of parenthood. Afterwards the midwife summed up the reflections of the participants and further commented on common changes and challenges in the relationship during pregnancy and after birth, and the importance of good communication. A short film about communication skills was shown. . The film was developed by the Danish Centre of Family Development and illustrated firstly inappropriate communication between a newborn couple and then more appropriate communication strategies. The film was followed by practical couple-communication exercises. Women attending the classes alone either paired up or conducted the exercise with the midwife. In short the exercise entailed one person listening actively and without interrupting while the other person described thoughts and feelings about a certain topic. Afterwards the person listening was encouraged to acknowledge what she/he heard before changing roles. This communication exercise was used throughout the entire NEWBORN program covering different topics e.g.

expectations of parenthood, labour, the relationship after birth etc. The aim was to try to understand the other person's perspective before trying to be understood. The importance of one's own childhood when becoming a parent was also a topic in this session. Participants were asked to think back to their own childhood, how they were raised, their parents' parenting style (e.g. warm and affectionate, strict etc.), traditions etc. Afterwards they were to consider things they would like to carry forward into their own parenting as well as things they might want to do differently. Finally, the couple began to discuss the topic using the communication technique they had been taught. The aim was to start a thought process. As there was not sufficient time for long in depth discussions participants were encouraged to carry on the discussions at home. During the first break participants who wished to do so were asked to write down their contact details so that they could be shared in the group. Throughout all the sessions the midwife had a facilitating role helping discussions along if needed and commenting where appropriate. At the end of all sessions participants were asked to consider and write down the most important take home points of the session.

Suggested preparation for the next session: seven short informational film clips (duration between 2 and 7 minutes) on the first signs of labour, the time at home in early labour, birth, when there is a need for obstetric intervention, and pain relief. The films, exercises, and written information on session topics were available on the network website.

Session 2 (approximately 33 weeks gestation):

• Expectations in relation to birth

• The normal course of labour

• When there is a need to intervene in labour

• Pain relief and coping strategies

• Partner support during labour

The aim of this session was to provide the participants with information, and enhancing their existing knowledge and understanding of the normal course of labour, pain-relief, and what might be expected if there is a need for obstetric intervention. First, the midwife gave a short verbal presentation about the advantages of considering one's expectations for the birth without adopting any firm success criteria. Next, couples discussed their hopes and expectations for labour and birth using the communication framework - they were asked to consider their individual resources and action competencies in relation to increasing the likelihood of obtaining their wishes. Also the couples were asked to discuss how they might support one another during birth and labour – using practical examples. For the topic on pain relief the women were first asked to discuss their thoughts and previous experiences with coping with pain and physical and mental strain – what did they do, what helped them, can they use any of these strategies during labour? Next they discussed their thoughts and knowledge on various methods of pain relief in a group with the other women. Meanwhile the men discussed their thoughts and feelings about their role during labour and birth. Plenary discussions and summing up thoughts and ideas were used so that participants could learn from and be inspired by one another. Vicarious learning and feedback were considered important in relation to self-efficacy.

Suggested preparation for the next session: participants were encouraged to ask women in their social network about their breastfeeding experiences, and read a pamphlet that was handed out on breastfeeding.

Session 3 (approximately 35 weeks gestation):

• Feeding a newborn –including breastfeeding intention, expectations, facts and myths

• Interpreting the newborn's signs, symptoms and behaviour

• Taking care of a newborn – baby cues, bonding, and practical issues

• Mood swings and postnatal depressive symptomatology

Participants discussed wishes for feeding their newborn and feeding experiences in their networks in small groups. The midwife then gave a short verbal presentation about how expectations, support, and the advice received from family and friends may affect e.g. breastfeeding intention and perseverance in the case of difficulties. Bearing the breastfeeding experiences of individuals in their social networks in mind (preparation for this session), participants were encouraged to consider who it might be most helpful to seek breastfeeding support and advice from if necessary.

Afterwards, the participants did a group exercise. Cards with a variety of breastfeeding topics were spread out on the table, and participants were asked to pick a topic that they wished to hear more about, and tell the group why they had picked the chosen topic. Topics included e.g. how to tell that the baby is getting enough milk, positioning, importance of partner support, feeding patterns, breast engorgement etc. There were certain topics that the midwife was told to cover regardless of whether it had been picked by a participant or not e.g. how to tell the baby is getting enough milk to ensure that participants was informed about vital topics. The pilot study showed that participants were likely to choose a topic they already knew something about in order to receive verification and feedback from the midwife and from the other participants (unpublished data) – this may help increase self-efficacy.

The midwife then gave information and showed short film clips on baby cues and sleep patterns. The participants were encouraged to give examples of the newborn's senses, and the importance of communicating with the newborn was underlined. Information on the prevention of cot death was given. Next the initial time at home with a newborn and the importance of social networks for emotional and practical support was discussed (the group was considered a potential supportive social network).

Participants were given an exercise where they were asked to fill in a list of expected daily activities after the baby was born. Afterwards they compared their list with their partner's or that of another group member before summing up in plenum. The aim of this exercise was to increase awareness of what changes life with a newborn has on a daily routine, how much time is spent on breastfeeding etc. Participants were also asked to consider activities that give them energy and pleasure (e.g. playing football, going out with friends, reading a book), and how they might incorporate some of these activities in their new daily lives. Next participants were encouraged to reflect upon how they normally handle worries, and to discuss this topic with their partner. Finally common emotional reactions and postnatal depressive symptomatology was covered. The importance of being open about these emotions and supportive of one another was stressed, as was the importance of seeking help when deemed necessary.

Session 4 (approximately 5 weeks post-partum):

• Birth experiences

• Mood swings and postnatal depressive symptomatology

• The first time at home with a newborn – experiences, challenges and solutions

• Couplehood – partner support, communication, division of household tasks

This session was facilitated by a midwife as well as a healthcare visitor. The aim was for the newborn parents to share birth experiences, and their experiences in their new roles as parents so far. The topic of common emotions and postnatal depressive symptomatology was revisited. Next groups of four were asked to discuss how being a parent was different to what they expected, which challenges they considered to be the greatest, how they coped with/handled these challenges, and what had been the greatest joys. The parents were able to ask the midwife and the healthcare visitor practical questions during the break.

After the break, using the communication framework, couples were asked to discuss what the best change has been in their partner after becoming a parent, what has worked really well in the relationship, and in sharing household tasks, and what could make it even better. Finally the healthcare visitor talked about sex (including contraception) and intimacy after becoming parents. Before leaving the session, the parents were encouraged to discuss if and how they would keep in contact with the group.

Table S1: Odds ratios (OR) (95% confidence interval (CI)) and relative risks (RR) (95% CI) for use of epidural analgesia in best case and worst case scenarios when comparing the experimental program to standard lectures.

	Experimental	Control	Adjusted OR (95% CI)*	P-value*	Crude OR (95% CI)	P-value	Crude RR (95% CI)	P-value
Best-case scenario	265/883 30.0%	278/883 31.5%	0.93 (0.76-1.14)	0.49	0.93 (0.76- 1.14)	0.50	0.95 (0.83-1.10)	0.50
Worst-case scenario	290/883 32.8%	248/883 28.1%	1.25 (1.02-1.54)	0.03	1.25 (1.02-1.53)	0.03	1.17 (1.02-1.34)	0.03

*Adjusted for trial stratification variables: vulnerability and parity.

Table S2: Use of birth and parent preparation programs in the intervention and control groups of the NEWBORN Trial (N=1545).

	Experimental	Control
Did not attend any birth and parent preparation	19/764 2.5%	87/781 11.1%
Attended the NEWBORN program	674/764 88.3%	2/781 0.3%
Attended the auditorium-based lecture on birth at Hvidovre Hospital	129/764 16.9%	551/781 70.6%
Attended the auditorium-based lecture on breast-feeding at Hvidovre Hospital	105/764 13.8%	488/781 62.5%
Attended other types of birth and parent preparation offers *	191/764 25.0%	302/781 38.7%

* This group contains a broad range of preparation offered by private stake holders, e.g., physical exercise classes, yoga classes, and mental preparation for delivery (called 'painless delivery').

Paper VI

Sexual & Reproductive Healthcare ■■ (2016) ■■–■■

Contents lists available at ScienceDirect

Sexual & Reproductive Healthcare

journal homepage: www.srhcjournal.org

Short communication

Antenatal education in small classes may increase childbirth self-efficacy: Results from a Danish randomised trial

Carina S. Brixval *, Solveig F. Axelsen, Lau C. Thygesen, Pernille Due, Vibeke Koushede

National Institute of Public Health, University of Southern Denmark, Øster Farigmagsgade 5A, 1353 Copenhagen C, Denmark

ARTICLE INFO

Article history:
Received 30 November 2015
Revised 22 February 2016
Accepted 14 March 2016

Keywords:
Antenatal education
Childbirth self-efficacy
Randomised trial
Pregnancy

ABSTRACT

Antenatal education in small classes may increase childbirth self-efficacy. In this randomised trial we assessed the effect of a structured antenatal programme versus auditorium-based lectures on childbirth self-efficacy measured by three single items. We found that women in the intervention group reported statistically significant higher levels of confidence in their ability to cope at home during labour compared to the control group. Likewise, the intervention had a positive effect on the women's confidence in own ability to handle the birth process.

Introduction

Childbirth self-efficacy reflects a woman's trust in her ability to cope with labour and birth. Childbirth self-efficacy may influence birth experience as well as obstetric interventions; e.g. a high level of childbirth self-efficacy is associated with lower levels of anxiety, pain, and obstetric intervention compared to low levels of childbirth self-efficacy [1].

According to the theory of self-efficacy there are four main sources of self-efficacy: (1) personal experiences, (2) vicarious experiences, e.g. by hearing about or observing other people's experiences, (3) social persuasion by encouragement from others, and (4) emotional interpretations of physical states [2].

Antenatal education in small classes may provide a suitable environment for enhancing expectant parents' self-efficacy. Discussing feelings and concerns related to birth and parenthood with a midwife and couples in a similar situation, may contribute to valuable social networks, enhance parents' awareness of their own resources, and increase their confidence in their ability to cope with the delivery.

To date, insufficient evidence exists as to whether antenatal education in small classes has an effect on obstetric or psycho-social outcomes [3]. The aim of the present short communication is to present the effect of a structured antenatal education programme in small classes versus auditorium-based lectures on childbirth self-efficacy.

Methods

Design

Data used in this short communication stem from the Danish randomised NEWBORN trial [4]. The effect of the programme on the primary and secondary outcome measures will be reported elsewhere. The outcome reported in this paper, childbirth self-efficacy, is one of the intermediate trial outcomes.

A detailed description of the trial is published in a design article [4]. A total of 1766 women from the largest birth site in the Capital Region of Denmark were randomised to either the intervention group (n = 883) or control group (n = 883) after informed consent was received. The intervention consisted of three sessions of antenatal education in small classes for the duration of 2.5 hours per session. The programme focused on strengthening relationships and improving information and problem solving skills for expectant parents in order to ease birth and the transition to parenthood.

The intervention programme intended to increase sources of self-efficacy e.g. through social modelling, support, and identification of coping strategies to reduce stress reactions and negative interpretations related to labour pain.

The control group received the standard care offer consisting of two auditorium-based antenatal lectures each lasting 2 hours.

Data collection

Data on childbirth self-efficacy was collected by web-based questionnaires in gestation week 37 and was measured by three single items developed for the NEWBORN trial; (1) I believe I will feel confident at home once labour has begun (e.g. before going to the labour

* Corresponding author. National Institute of Public Health, University of Southern Denmark, Øster Farigmagsgade 5A, 1353 Copenhagen C, Denmark. Tel.: +45 65507845.
E-mail address: cabr@niph.dk (C.S. Brixval).

http://dx.doi.org/10.1016/j.srhc.2016.03.003
1877-5756/

ward), (2) I believe that I can contribute to making the birth a good experience, and (3) I believe that I will be able to handle the birth process no matter how it turns out. All items had the following response categories: totally agree, agree, neither/nor, disagree, and totally disagree. For analysis responses were trichotomised into; high self-efficacy (totally agree, agree), neither/nor, and low self-efficacy (disagree and totally disagree).

Of the 1766 randomised women, 1508 (85%) women returned the questionnaire. Of these 165 women had already given birth and were excluded from analysis. For each childbirth self-efficacy item only between four and eight individuals had missing answers.

Data analysis

Data were analysed according to the intention-to-treat principle. Differences in childbirth self-efficacy between the intervention group and control group were tested in multinomial logistic regression models adjusted for the protocol specified stratification variables, parity (primiparous or multiparous) and vulnerability (yes or no as defined by their general practitioner). There were eight criteria listed for vulnerability, for example; former or current psychiatric disorder, adverse psycho-social background, or concerns about parenting skills. Odds ratios (OR) and 95% confidence intervals (95% CI) were calculated. High childbirth self-efficacy was used as reference category.

Information on childbirth self-efficacy was provided by 75.6% of the women in the intervention group and 77.3% of the women in the control group. We used Inverse Probability Weighting to account for the potential bias related to the missing values [5].

Analyses were performed using SAS v. 9.3, SAS Institute Inc.

The NEWBORN trial has been assessed and registered by the Capital Region's ethics committee (CVR/SE-nr: 30113713), in the Danish Data Protection Agency (reference number: 2011-54-1289), and at ClinicalTrials.gov (ID: NCT01672437).

Results

In the intervention group, fewer women (4.1%) felt low confidence in their ability to cope at home during labour compared with the control group (8.0%) (Table 1). When examining differences between groups in multinomial logistic regression models, the adjusted odds ratio for low self-efficacy was 0.48 (95% CI: 0.32–0.73, $p < 0.001$) in the intervention group compared with the control group. No significant difference was seen between intervention groups when comparing the category "neither/nor" with high self-efficacy. Only a total of five women felt low confidence in own ability to make the delivery a positive experience and there were no statistically significant differences between the two intervention groups. Fewer women in the intervention group (5.0%) felt low confidence in own ability to handle the birth process compared with the control group (7.4%). The adjusted odds ratio for low self-efficacy was 0.66 (95% CI: 0.44–0.98, $p = 0.04$) in the intervention group compared with the control group. No significant difference was seen between intervention groups when comparing the category "neither/nor" with high self-efficacy.

Discussion

The results from the NEWBORN trial indicate that attending a structured antenatal education programme in small classes may increase confidence in own ability to cope at home during labour and confidence in own ability to handle the birth process. A small proportion (8% or below) had low childbirth self-efficacy measured by the three items. To our knowledge, there exist no trials comparable to the NEWBORN trial that assess the effect of small classes on

Table 1
Distribution of measures of childbirth self-efficacy in intervention and control groups. Odds ratios (OR) (95% CI) for low and neither/nor versus high childbirth self-efficacy in the intervention group compared with the control group.

	High self-efficacy % (N)	Neither/nor % (N)	Low self-efficacy % (N)	Low versus high			Neither/nor versus high		
				Unadjusted odds ratio (95% CI)	p-value	Adjusted odds ratio (95% CI)[a]	p-value[a]	Unadjusted odds ratio (95% CI)	p-value	Adjusted odds ratio (95% CI)[a]	p-value[a]
Confidence in ability to cope at home during labour											
Intervention	78.4 % (519)	17.5 % (116)	4.1% (27)	0.48 (0.32–0.73)	<0.001*	0.48 (0.32–0.73)	<0.001*	1.04 (0.81–1.33)	0.78	1.04 (0.81–1.33)	0.79
Control	75.8% (513)	16.3% (110)	8.0% (54)								
Confidence in own ability to make the delivery a positive experience											
Intervention	93.9% (620)	5.8% (38)	0.3% (2)	0.66 (0.14–3.10)	0.60	0.67 (0.14–3.16)	0.62	0.72 (0.50–1.05)	0.09	0.72 (0.50–1.05)	0.09
Control	91.7% (619)	7.9% (53)	0.4% (3)								
Confidence in own ability to handle the birth process no matter how it turns out											
Intervention	68.8% (455)	26.2% (173)	5.0% (33)	0.65 (0.44–0.97)	0.03*	0.66 (0.46–0.98)	0.04*	1.02 (0.83–1.27)	0.83	1.03 (0.83–1.28)	0.80
Control	67.7% (458)	25.0% (169)	7.4% (50)								

[a] Adjusted for trial stratification variables, vulnerability and parity.
* $p < 0.05$.

Please cite this article in press as: Carina S. Brixval, Solveig F. Axelsen, Lau C. Thygesen, Pernille Due, Vibeke Koushede, Antenatal education in small classes may increase childbirth self-efficacy: Results from a Danish randomised trial, Sexual & Reproductive Healthcare (2016), doi: 10.1016/j.srhc.2016.03.003

childbirth self-efficacy [3]. Hence, we have no former studies to compare the findings with.

We used three single items developed for the NEWBORN trial as a proxy for childbirth self-efficacy [4]. The items cover essential elements of childbirth self-efficacy, i.e. confidence in own ability to cope with labour in the latent phase of labour and ability to cope with the birth process. However, it is possible that the measures are too crude to capture the complexity of the self-efficacy concept. A comprehensive 62-item scale for measuring childbirth self-efficacy has been developed: the Childbirth Self-Efficacy Inventory (CBSEI) [6]. Although measuring childbirth self-efficacy by this scale might have contributed with more thorough details about the women's self-efficacy, it was not feasible to include the long scale in the questionnaires. However, we believe that the single item questions are good indicators of the women's childbirth self-efficacy because they capture the essential elements in Bandura's self-efficacy theory; confidence in own ability to perform specific behaviours [2].

Women with low childbirth self-efficacy may cost more e.g. through early admissions and pain relief, making a reduction in the proportion of women with low childbirth self-efficacy a relevant public health issue. The NEWBORN trial was carried out among a population of well-educated women. It is possible that the proportion of low childbirth self-efficacy as well as the intervention effect is different among other population groups. This needs further investigation.

Acknowledgements

This work was supported by the Danish Cancer Society. The funders have not been involved in trial design, data collection and analysis, or decision to publish this manuscript. The authors have no other conflicts of interest to disclose.

References

[1] Carlsson IM, Ziegert K, Nissen E. The relationship between childbirth self-efficacy and aspects of well-being, birth interventions and birth outcomes. Midwifery 2015;31(10):1000–7.
[2] Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977;84(2):191–215.
[3] Brixval CS, Axelsen SF, Lauemøller SG, Andersen SK, Due P, Koushede V. The effect of antenatal education in small classes on obstetric and psycho-social outcomes-a systematic review. Syst Rev 2015;4(1):1–9.
[4] Koushede V, Brixval CS, Axelsen SF, Lindschou J, Winkel P, Maimburg RD, et al. Group-based antenatal birth and parent preparation for improving birth outcomes and parenting resources: study protocol for a randomised trial. Sex Reprod Healthc 2013;4(3):121–6.
[5] Seaman SR, White IR. Review of inverse probability weighting for dealing with missing data. Stat Methods Med Res 2013;22(3):278–95.
[6] Lowe NK. Maternal confidence for labor: development of the Childbirth Self-Efficacy Inventory. Res Nurs Health 1993;16(2):141–9.

Please cite this article in press as: Carina S. Brixval, Solveig F. Axelsen, Lau C. Thygesen, Pernille Due, Vibeke Koushede, Antenatal education in small classes may increase childbirth self-efficacy: Results from a Danish randomised trial, Sexual & Reproductive Healthcare (2016), doi: 10.1016/j.srhc.2016.03.003