DUTCH PERINATAL SYSTEM

PERFORMANCE

AND

INNOVATIVE STRATEGIES

J. van der Kooy
2013

Dutch Perinatal System: Performance and Innovative Stategies

PhD thesis, Erasmus University Rotterdam, The Netherlands

The research presented in this dissertation was performed at the department of Obstetrics and Gynaecology, division of Obstetrics and Prenatal Medicine, Erasmus University Medical Centre, Rotterdam, The Netherlands

The printing of this thesis was financially supported by:

The department of Obstetrics and Gynaecology, Erasmus University Medical Centre

The Rotterdam municipal authorities and the 'Ready for a Baby' programme (www.klaarvooreenkind.nl)

J.E. Jurriaanse Stichting

Medical Dynamics

BMA-MOSOS

Goodlife Pharma

Ferring B.V.

ChipSoft

Bayer HealthCare

MicroWare Automatisering – Onatal

Heinen & Löwenstein

Smith & Nephew Nederland C.V.

ISBN/EAN 978-94-91487-07-1

Layout 358atwork

Print Medix Publishers

Cover design: Maria van Oosten

Layout and printing: Medix Publishers

Dutch Perinatal System

performance en innovative strategies

Het Nederlands verloskundig systeem
performance en innovatieve strategieën

PROEFSCHRIFT

ter verkrijging van de graad van doctor aan de

Erasmus Universiteit Rotterdam

op gezag van de

rector magnificus

prof.dr. H.G. Schmidt

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

18 september 2013 om 13.30 uur

door

Jacoba van der Kooy

geboren te Delft

PROMOTIECOMMISSIE

Promotoren	Prof. dr. E.A.P. Steegers
	Prof. dr. G.J. Bonsel
Overige leden	Prof. dr. A.J.P. Schrijvers
	Prof. dr. I.K.M. Reiss
	Prof. dr. R.J. Stolker
Copromotoren	Dr. S. Denktaş
	Dr. E. Birnie

Paranimfen	Virender Jashvant John Poeran
	Florina Wilhelmina Gerestein – van der Kooy

For Him and through Him and to Him are all things.

To Him be glory forever! Amen. (Romans 11:36)

CONTENTS

Chapter 1 Introduction **1**

Part I What are possible explanations for the underperformance of the Dutch perinatal care system, in particular focussing on risk selection and midwife-led birth care?

Chapter 2 The effectiveness of risk selection in the Dutch obstetric care system **15**

Chapter 3 Planned home compared with planned hospital births in The Netherlands: intrapartum and early neonatal death in low-risk pregnancies **33**

Obstetrics and Gynecology. 2011;118:1037-46.

Chapter 4 Planned home compared with planned hospital births in the Netherlands: intervention rates and intervention specific mortality rates in low-risk pregnancies **53**

Part II How can Dutch perinatal care be improved, in particular, are there innovative strategies available to improve risk selection and midwife-led birth care?

Chapter 5 Different settings of place of midwife-led birth: evaluation of a midwife-led birth centre **75**

Chapter 6 Improving the prediction of small-for-gestational-age (SGA) in an assumed low-risk population **95**

Chapter 7 A newly developed scavenging system for administration of nitrous oxide during labour: Safe occupational use **119**

Acta Anaesthesiologica Scandinavica. 2012:920-5.

Chapter 8 Evaluation of the use of Nitrous Oxide analgesia during labor in the midwife-led birth centre Sophia **129**

Part III How can the performance of the Dutch perinatal care be evaluated on non clinical aspects of quality of care?

Chapter 9 Validity of the ReproQ: a questionnaire measuring the World Health Organization concept of Responsiveness in Perinatal Care **151**

Chapter 10 The relative importance of non-clinical quality of care domains within the Dutch perinatal system **175**

Chapter 11 Measuring the WHO concept of Responsiveness in Perinatal Care: Outcomes of the ReproQ questionnaire **191**

Chapter 12 General discussion **217**

Chapter 13 Summary/samenvatting **231**

Chapter 14 Authors and Afflictions **241**

Portfolio **243**

Word of Thanks/dankwoord **246**

Addendum Differential susceptibility to early literacy intervention in children with mild perinatal adversities: short and long-term effects of a randomized control trial **249**

Journal of Educational Psychology, 2012:p337-349

1

Introduction

INTRODUCTION

The Dutch perinatal care system

The organization of the Dutch perinatal care system is unique since, in contrast to most other high-income countries, certified community midwives play an independent role in the care for pregnant women. Community midwives care for pregnant women with an assumed or estimated low-risk for medical complications starting from the early prenatal until the postpartum period. If during pregnancy no risks are detected, women have the option of delivering at home, at a birthing centre or in a hospital, in all cases under the supervision of their community midwife. Should complications (threaten to) occur, community midwives refer women under their supervision to secondary care by obstetricians in a hospital setting. If necessary, secondary caregivers then refer women who are severely ill and/or have threatened pregnancies to tertiary perinatal care, which is located in academic hospitals and in non-academic hospitals with obstetric high care and neonatal intensive care units. Approximately one out five women directly starts antenatal care at a secondary or tertiary hospital due to their initial high-risk status.

The functioning of this unique system depends on the mutual cooperation of the health care professionals involved, the availability of (different) facilities, the absence of financial barriers, and adequate information to the patients.[1]

Underperformance of the Dutch perinatal care system

The merits of this system have come under scrutiny since the national perinatal mortality rate showed to be one of the highest in Europe.[2-4] Other concerns about the Dutch perinatal system include economic and process inefficiency of the current risk selection, since high referral rates to secondary care exist during antepartum care and delivery. Concerns were also raised on the higher rate of medical interventions in planned hospital births compared with planned homebirths in assumed low risk women and concerns were raised on the negative perceptions of birth experiences with obstetric care.[5-7] A final concern is on the substantial maternal and perinatal health inequalities between ethnic minority groups, between low and high socio-economic class, and between urban and non-urban regions.[8-9]

Response to the underperformance of the Dutch perinatal care system

Over time, health care professionals, insurance companies, patient organizations and policy makers have initiated various improvements of Dutch perinatal health

on local and national level. Each actor addressed different aspects focusing either on health care factors, economic factors, patient factors, or environmental factors, resulting in a fragmented and often self interest guided response, rather than an integral response.

In 2009 the underperformance of the Dutch perinatal care system was further demonstrated in the so called 'Signalement-studie', a national report on the priority setting of perinatal research on behalf of ZonMw (The Netherlands Organisation for Health Research and Development).[10] The 'Signalement-studie' adopted a public health, system-oriented approach. It related the Dutch perinatal care system to patient, environmental and health care factors in an integrated way. The study introduced the concept of Big4; four perinatal conditions which precede perinatal mortality in 85% of cases. At least three of them offer some interventional opportunities in terms of diagnosis, preventive actions or treatments when these are detected timely (see further). The occurrence of these conditions and their progression in perinatal death are influenced by the accumulation and interaction of patient, environmental and health care factors.

At the same time as the issuing of the Signalement study, an advisory group on pregnancy and birth commissioned by the Dutch health authorities proposed a report in 2010, "A Good Beginning" (*Een Goed Begin*).[11] This rapport discussed possible causes and mechanisms underlying the poor performance of the Dutch perinatal care system, resulting in recommendations at the system level.

The preceding integral analysis and the suggestions to improve care provision was put into practice in the Rotterdam urban perinatal health program called 'Ready for a Baby'.[12] This program was instituted by the municipal council of Rotterdam and the Erasmus Medical Centre in 2009, as perinatal health in Rotterdam was even worse compared to the current national average. The main objective of this 10-year program was to improve perinatal health and to reduce perinatal mortality in all districts of Rotterdam to at least the current national average. It covered the five phases of perinatal health care namely; preconception-, antenatal-, birth-, post partum- and youth care phase and mainly focused on improving risk selection and structural cooperation across disciplines. The five different phases were connected to each other both on the clinical and on the organizational level ('ketenzorg') in agreement with the '"Good Beginning" report.[11] Programs such as 'Ready for a Baby' and numerous other initiatives directed to perinatal health improvement, all require methods to detect deficits, to guide the changes in process and outcome and to assess the final effects of these changes. While some of these methods were already available, new ones had to be developed.

AIMS OF THIS THESIS

This thesis focusses on the possible explanations for the relative high Dutch perinatal mortality, on innovations put forward for improvement, and on the evaluation methods required to proof the benefits claimed. We introduce the concept of Big4 as a new method to improve case mix adjustment in observational studies and to analyse the occurrence of perinatal mortality in two steps: the development towards perinatal morbidity, and the subsequent occurence of mortality. It further elaborates on methods to assess the non clinical aspects of quality of care, introducing the responsiveness concept in perinatal care. The general responsiveness concept was introduced by the World Health Organization (WHO), and is defined as the way an individual is treated and the environment in which she is treated during health system interactions.[13] As universal concept used on a global scale and applied to health systems, it seemed - after tailoring to perinatal care - suitable to measure the chain of care as experienced by the heterogeneous Rotterdam population.

This thesis consists of three questions and ten subquestions:

Part I. What are possible explanations for the underperformance of the Dutch perinatal care system, in particular focussing on risk selection and midwife-led birth care?

1. How does the current risk selection of the Dutch perinatal care system perform? (chapter 2)
2. Does planned place of birth in community midwife-led deliveries influence perinatal mortality outcomes? (chapter 3)
3. Does planned place of birth within community midwife-led deliveries influence intervention rates and, subsequently, perinatal mortality? (chapter 4)

Part II. How can Dutch perinatal care be improved, in particular, are there innovative strategies available to improve risk selection and midwife-led birth care?

4. Does the introduction of a midwife-led birth centre improve perinatal outcomes and the indication for interventions compared to other planned places of birth? (chapter 5)
5. Is the prior detection of Small for Gestational Age improved when an adjuvant risk selection tool at birth is introduced? (chapter 6)
6. Can occupational safety be reached in the midwife-led birth centre Sophia when reintroducing nitrous oxide analgesia during labour? (chapter 7)
7. What is the effect of the introduction of nitrous oxide analgesia in a midwife-led birth centre on patient flow and pain relief? (chapter 8)

Part III. How can the performance of the Dutch perinatal care be evaluated on non clinical aspects of quality of care?

8. What are the psychometric properties of the newly developed responsiveness questionnaire ReproQ? (chapter 9)
9. Should different domain weight adjustment be used for subgroups within the Dutch perinatal care system? (chapter 10)
10. What are the responsiveness outcomes of maternal experiences during perinatal care using the ReproQ? (chapter 11)

Part I. What are possible explanations for the underperformance of the Dutch perinatal care system, in particular focussing on risk selection and midwife-led birth care?

Challenges arise when evaluating the Dutch perinatal care system since the primary outcome perinatal mortality is rare and since its causes are multifactorial, consisting of patient-, environment- and health-care related factors, which may occur during the entire gestational period which starts before conception. In addition, the Dutch perinatal care system is organized in a complex way. It consists of primary, secondary, and tertiary care; referral may happen during pregnancy, delivery and thereafter, and care is provided at different places including midwife-led birth at home, in birthing centres or a hospital.

Pivotal to the Dutch perinatal system is the estimation of a pregnant woman's risk status (in comparative analysis pointed to as case mix), since risk status determines obstetric care, the professional(s) involved and the place of birth. The community midwife offers care to low and medium risk women, and the gynaecologist to high risk women. Assuming risk selection is effective at least to an acceptable degree, the case mix of pregnant women attending primary and secondary care should be rather different when the effectiveness of risk-led care is evaluated, and few high risk cases are left under midwife's supervision alone. In contrast, case mix is assumed to be equal among low-risk women attended in primary care when different modalities of midwife-led birth care (home vs. hospital vs. birth centre) are compared. To evaluate the effectiveness of risk selection and the health care performance at different places of birth, detailed risk information when making comparisons is needed, since both can only be assessed within observational constraints.[14-21] However, in most national registries, including ours, detailed risk information (case mix) is unavailable, or, for some factors, partially as recording is not obligatory.

In order to gain more insight in the causes of perinatal mortality and in the single or joint impact of patient-, environment-, and health care related factors, the 'Signalement study' introduced the concept of Big4.[10] This concept of Big4 was attained from detailed analysis of the complete perinatal dataset of the Netherlands Perinatal Registry (PRN), years 2000-2007, containing 1.25 million records of pregnancies.[10] The analysis showed that any of four specific well defined conditions precede perinatal mortality in 85% of cases. These conditions were defined as: (1) congenital abnormalities (list defined); (2) intrauterine growth restriction (SGA, birth weight below the 10th percentile by gestational age, gender and parity); (3) preterm birth (< 37th week of gestation); or (4) low Apgar score (< 7, measured at 5 minutes after birth). Factors influencing the aetiology of Big4 were distinguished from factors influencing its course ('prognosis') once a Big4 exists. The resulting two-staged model gives more insight in causes leading to perinatal mortality, hereby creating insight in opportunities for corresponding intervention strategies (see figure 1.1). It hypothesised that an accumulation of risk factors, consisting of patient, environment and health care related factors, leads to the occurrence of Big4 (etiology), and subsequently influences its prognosis. The prevalence of Big4 then can serve as a proxy of the population's risk status (case mix) and thus be used in two ways. (1) Within the analyses of the quality of the Dutch perinatal risk selection system, the Big4 (with the exception perhaps of asphyxia, low Apgar) may represent the outcome measure used to evaluate succes of antenatal risk selection. (2) When

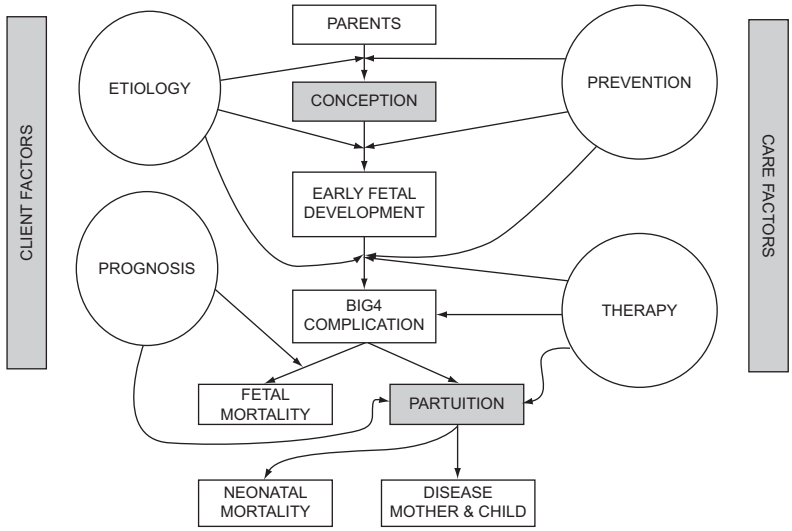


Figure 1.1 Two-staged model of the various causes, including client factors and care factors, leading to the occurrence of a Big4 complication and causes eventually leading to fetal and neonatal mortality. (Bonsel GJ et al, Lijnen in de Perinatale Sterfte, Signalementstudie Zwangerschap en Geboorte 2010)

the quality of midwife-led birth care in different settings is compared, the Big4 represents an objective estimate of the risk load at birth, suitable as adjustment factor.

Part II. How can Dutch perinatal care be improved, in particular are there innovative strategies available to improve risk selection and midwife-led birth care?

Strategies to improve the quality of care can be arranged according to Donabedian's model. This model was designed to analyse the quality of care of health care systems and includes three levels: structure, process and outcome.[22-23] Structure refers to prerequisites, such as setting, staff and equipment. Process describes the professional activities associated with care. Outcome refers to evaluation of goal achievements. An innovation strategy can affect all three levels of Donabedian's model. Strategies to improve the Dutch perinatal care system on the structure level include: more integral organization of care (shared care), centralization of delivery care, continuous availability of acute care with senior staffing, introduction of midwife-led birth settings, connecting the different perinatal phases on an organizational level ('ketenzorg'), and a different dividing of financial funding and insurance. Strategies on the general process level include: closer co-operation between health care professionals, improving risk selection and risk management, the introduction or improved use of (new) diagnostic tools and therapies, the transition to a more proactive attitude towards pregnancy, the transition of care provision in a more patient oriented way. Strategies on the outcome level include the application of specific interventions in special cases e.g. the specific use of (new) therapies in specific patient groups.

One innovation strategy to improve Dutch perinatal care is the introduction of a new birth setting, namely a midwife-led birth centre. Its introduced changes at all levels. In our thesis it served as an important place to implement and evaluate innovations.

The birth centre Sophia, an midwifery care unit adjacent to the hospital, was established in October 2009.[24] It was designed to provide an intermediate setting of care between home and hospital birth for midwife-led birth care for low risk women. It is a separate and independently functioning unit, of four single birthing rooms, located on the same floor as the obstetric labour ward (100 meters). The unit is staffed by local community midwives. Low risk women, as defined by Dutch guidelines, may choose to deliver at the birth centre Sophia under supervision of their own community midwife.[25] These midwives, and the maternity care provider are co-responsible, together with the hospital, for the facility as business unit.

Table 1.1 gives an overview of the innovations that are currently implemented in the birth centre Sophia. The general strategy of the unit is to provide a safe, homelike environment

Table 1.1 An overview of possible innovations strategies on the different levels and the strategies that are currently implemented in the birth centre Sophia

Level	Definition	Examples	Birth Centre Sophia	This thesis
Structure				
	Prerequisites; setting, staff and equipment	Integral organization of care (shared care) Centralization of delivery care Continuous availability of acute care	Providing a safe, homelike environment adjacent to a hospital	X
		Introduction of new birth settings Connecting the perinatal phases on an organizational level ('ketenzorg')	Providing care during labour and post partum care Face to face referral to youth care in case of high risk family's	
Process				
	Professional activity associated with care	Closer co-operation between health care professionals	An expert group, consisting out of a gynaecologist, midwives. maternity nurses, public health experts and administrating staff is continually responsible for provided care and innovations Assessing the risk status of each women and act accordantly using standardised protocols	X
		Improving risk led care	Special attention is given to ethnic groups and women with a low social economic background	
		Patient oriented care Use of (new) diagnostic tools Use of (new) therapies Proactive attitude towards pregnancy ('culture')		
Outcome				
	Patient specific innovations	Use of patient specific (new) diagnostic tools Use of patient specific (new) therapies	Reintroduction and evaluation of Nitrous Oxide analgesia during labour	X

where every woman can give labour and receive postpartum care (while different perinatal phases are connected). The unit aims to provide risk led care, including assessing the risk status of each patient at different time intervals and act accordantly using standardised protocols (*improvement of risk led care*). Special attention is given to certain ethnic groups and women with a low social economic background (*patient centred care*). Protocols, new interventions and adverse outcomes are continuously monitored by an expert group consisting of a gynaecologist, midwives, maternity nurses, a public health expert and administrating staff (*integral organization of care, closer co-operation between health care professionals*). It also seeks to develop new interventions to improve midwife-led birth care (*use of new therapies*).

In this thesis we elaborate on the assessment of the midwife-led birth centre as a new birth setting focusing on clinical outcomes and possible use of a new therapy, being the reintroduction of nitrous oxide as an analgesic option during labour. In addition, we assess innovation strategies improving risk led care.

Part III. How can the performance of the Dutch perinatal care be evaluated on non clinical aspects of quality of care?

In 1948 the World Health Organization (WHO) defined health as 'a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity'. It was recognized that health systems must address the medical needs of individuals, but that it should also focus on the non clinical aspects affecting their well being.[13]

The measurement of non clinical aspects resulted in questionnaires on patient satisfaction. Patient satisfaction tries to capture patient perceptions of the quality of care and is thought to be an interaction among patient's preferences, patient's expectations, and patient's actual experience.[26-27] Therefore, difficulties may arise since patient satisfaction surveys may not capture what actually happens when patients come in contact with the health care system, since satisfaction is strongly influenced by his/hers expectations and characteristics.[27] Accordingly, literature showed that expectations do vary across individuals and populations, both between and within countries.[28-29] Other difficulties include the type and the components of interaction. Patient satisfaction surveys often capture general attitudes or satisfaction over longer periods. In addition, patient satisfaction surveys generally focus upon interactions in medical facilities alone, and often patient satisfaction surveys generally include both clinical and non clinical components of an interaction, where the clinical adequacy (the choice of diagnostic or treatment modality) may be impossible to judge for the client.[13]

Partly for these reasons, focus has moved from patient satisfaction surveys to surveys and interviews capturing more factually patients actual experiences with the health system. Responsiveness is one of the available approaches to address actual experiences of the non clinical aspects of health care, introduced by the World Health Organization (WHO). The WHO defines responsiveness as the way an individual is treated and the environment in which she is treated during health system interactions.[13] It contains non-financial, non clinical qualities of care ('domains') that reflect respect for human dignity and interpersonal aspects of the care process. The patient preferably reports on his or her own most recent experiences with the health system, reflecting on specified topics within these domains. The eight domains distinguished are: Dignity, Autonomy, Confidentiality, Communication, Prompt Attention, Social Consideration, Quality of Basic Amenities, and Choice and Continuity.

Responsiveness outcomes are already obtained for the whole Dutch health system.[30] However, a problem with the available WHO questionnaire is its general scope. Research and application by health insurance companies has therefore moved in recent years to more specific surveys with specific fields of interest.[26] Sofar responsiveness outcomes are unavailable for the Dutch perinatal care system, while such data are needed in view of the challenged quality of the system and the many innovations started. We developed the ReproQ, focussing on the Dutch perinatal care within the Netherlands. For integral judgement, the questionnaire covers the different disciplines and settings; it does not assume fixed combinations of disciplines, settings and client characteristics as the system changes rapidly in this respect. We tested the newly developed ReproQ for psychometric properties such as validity and general acceptability (including women with low socio economic status and a migrant background), As part of that we checked whether all domains where about equally valued by different subgroups.

REFERENCES

1. Posthumus AG, Scholmerich VL, Waelput AJ, et al. Bridging Between Professionals in Perinatal Care: Towards Shared Care in The Netherlands. Matern Child Health J. 2012 Dec 11.
2. Ravelli AC, Tromp M, van Huis M, et al. Decreasing perinatal mortality in The Netherlands, 2000-2006: a record linkage study. J Epidemiol Community Health. 2009 Sep;63(9):761-5.
3. Mohangoo AD, Buitendijk SE, Hukkelhoven CW, et al. [Higher perinatal mortality in The Netherlands than in other European countries: the Peristat-II study]. Ned Tijdschr Geneeskd. 2008 Dec 13;152(50):2718-27.
4. Zeitlin J, Wildman K, Breart G, et al. PERISTAT: indicators for monitoring and evaluating perinatal health in Europe. Eur J Public Health. 2003 Sep;13(3 Suppl):29-37.
5. Christiaens W, Gouwy A, Bracke P. Does a referral from home to hospital affect satisfaction with childbirth? A cross-national comparison. BMC Health Serv Res. 2007;7:109.
6. Rijnders M, Baston H, Schonbeck Y, et al. Perinatal factors related to negative or positive recall of birth experience in women 3 years postpartum in the Netherlands. Birth. 2008 Jun;35(2):107-16.

7. Amelink-Verburg MP, Rijnders ME, Buitendijk SE. A trend analysis in referrals during pregnancy and labour in Dutch midwifery care 1988-2004. BJOG. 2009 Jun;116(7):923-32.

8. Poeran J, Denktas S, Birnie E, Bonsel GJ, Steegers EA. Urban perinatal health inequalities. J Matern Fetal Neonatal Med. Apr;24(4):643-6.

9. de Graaf JP, Ravelli AC, Wildschut HI, et al. [Perinatal outcomes in the four largest cities and in deprived neighbourhoods in The Netherlands]. Ned Tijdschr Geneeskd. 2008 Dec 13;152(50):2734-40.

10. Bonsel GJ, Birnie E, Denktas S, Poeran J, Steegers EAP. Dutch report: Lijnen in de Perinatale Sterfte, Signalementstudie Zwangerschap en Geboorte 2010 (http://www.nvk.nl/Nieuws/Dossiers/DossierPerinatalezorg.aspx). Rotterdam: Erasmus MC; 2010.

11. van der Velden J. Een goed begin. Veilige zorg rond zwangerschap en geboorte. http://wwwknovnl/docs/uploads/Advies_stg_zw_en_geb_Een_goed_beginpdf_spread-versie_voor_mediapdf. 2009.

12. Denktas S, Bonsel GJ, Van der Weg EJ, et al. An urban perinatal health programme of strategies to improve perinatal health. Matern Child Health J. 2012 Nov;16(8):1553-8.

13. Valentine NB, de Silva A, Kawabata K, Darby C, Murray CJL, Evans DB (eds). Health system responsiveness: concepts, domains and measurement. Geneva: World Health Organisation; 2003.

14. Ackermann-Liebrich U, Voegeli T, Gunter-Witt K, et al. Home versus hospital deliveries: follow up study of matched pairs for procedures and outcome. Zurich Study Team. Bmj. 1996 Nov 23; 313(7068):1313-8.

15. Hildingsson IM, Lindgren HE, Haglund B, Radestad IJ. Characteristics of women giving birth at home in Sweden: a national register study. Am J Obstet Gynecol. 2006 Nov;195(5):1366-72.

16. Hogberg U. Homebirths in a modern setting - a cautionary tale. Acta Obstet Gynecol Scand. 2008;87(8):797-9.

17. Hutton EK, Reitsma AH, Kaufman K. Outcomes associated with planned home and planned hospital births in low-risk women attended by midwives in Ontario, Canada, 2003-2006: a retrospective cohort study. Birth. 2009 Sep;36(3):180-9.

18. Johnson KC, Daviss BA. Outcomes of planned home births with certified professional midwives: large prospective study in North America. Bmj. 2005 Jun 18;330(7505):1416.

19. Neuhaus W, Piroth C, Kiencke P, Gohring UJ, Mallman P. A psychosocial analysis of women planning birth outside hospital. J Obstet Gynaecol. 2002 Mar;22(2):143-9.

20. Wiegers TA, Keirse MJ, van der Zee J, Berghs GA. Outcome of planned home and planned hospital births in low risk pregnancies: prospective study in midwifery practices in The Netherlands. Bmj. 1996 Nov 23;313(7068):1309-13.

21. Hendrix M, Van Horck M, Moreta D, et al. Why women do not accept randomisation for place of birth: feasibility of a RCT in The Netherlands. Bjog. 2009 Mar;116(4):537-42; discussion 42-4.

22. Donabedian A. Explorations in quality assessment and monitoring: the definition of quality and approaches to assessment. Ann Arbor, Health Administration Press 1980.

23. Donabedian A. Evaluating the quality of medical care. Milbank Mem Fund Q. 1966 Jul;44(3):Suppl:166-206.

24. de graaf JP, Stam-Happel, M. , Schuur, O. , Willems, M.C., Wieren, W.J., Steegers, E.A.P. Bevallen in een geboortecentrum. Medisch Contact. 2003.

25. CVZ. CV. List of Obstetric Indications (in Dutch: Verloskundig Vademecum) 2003. Diemen (http://www.knov.nl/docs/uploads/Verloskundig_Vademecum_2003.pdf). 2003.

26. Sitzia J, Wood N. Patient satisfaction: a review of issues and concepts. Soc Sci Med. 1997 Dec;45(12):1829-43.

27. Ware JE, snyder MK, Wright WR, Davies AR (eds). Defining and measuring patient satisfaction with medical care. Evaluation and Program Planning. Volume 6: 247-263; 1983.

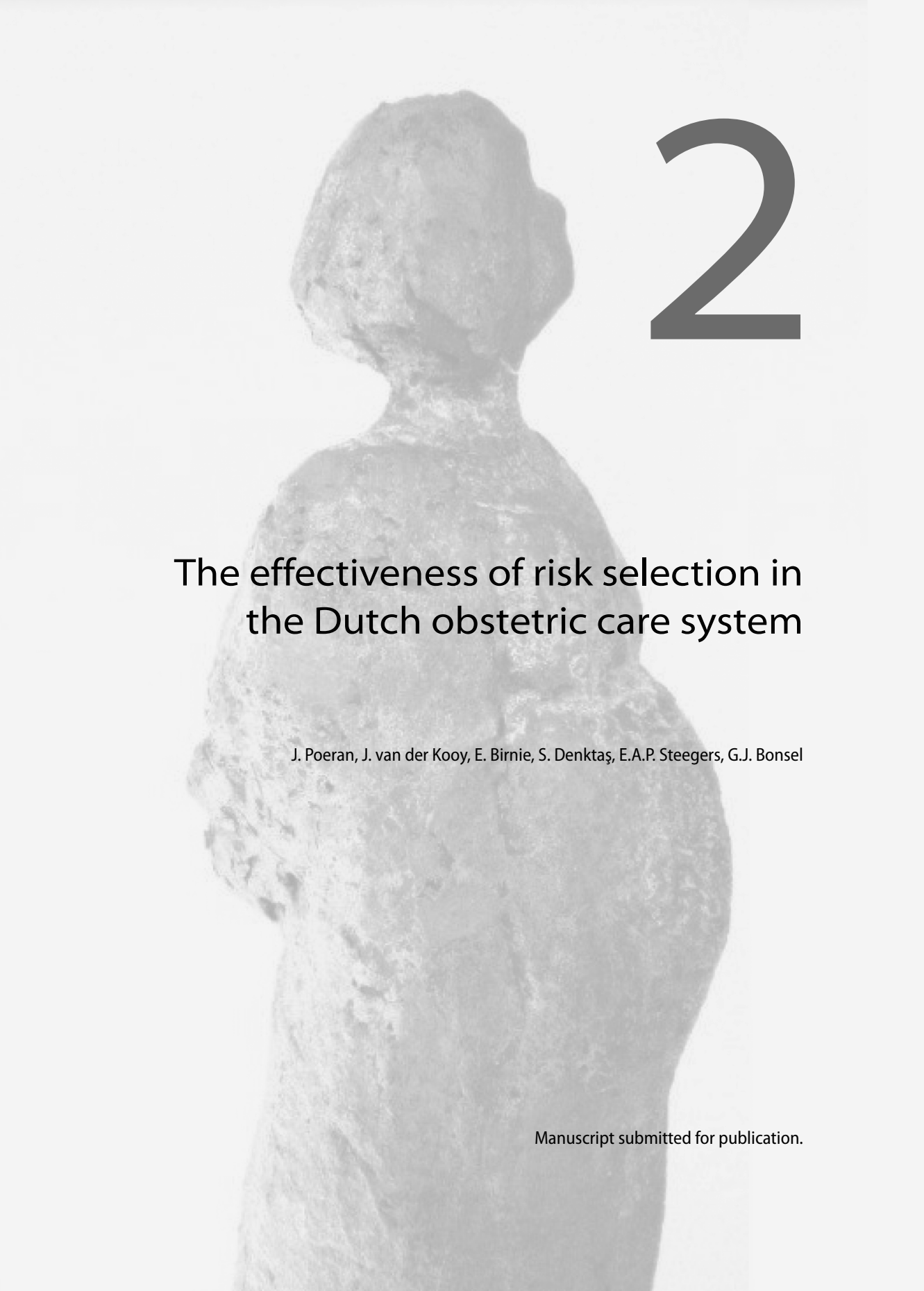
28. Blendon RJ, Kim M, Benson JM. The public versus the World Health Organization on health system performance. Health Aff (Millwood). 2001 May-Jun;20(3):10-20.

29. Murray CJ, Kawabata K, Valentine N. People's experience versus people's expectations. Health Aff (Millwood). 2001 May-Jun;20(3):21-4.

30. Valentine NB, de Silva A, Murray CJL. Estimating Responsiveness Level and Distribution for 191 Countries: Methods and Results. World Health Organization GPE Discussion Paper Series: No 22. 2000.

PART I

What are possible explanations for the underperformance of the Dutch perinatal care system, in particular focussing on risk selection and midwife-led birth care?



2

The effectiveness of risk selection in the Dutch obstetric care system

J. Poeran, J. van der Kooy, E. Birnie, S. Denktaş, E.A.P. Steegers, G.J. Bonsel

Manuscript submitted for publication.

ABSTRACT

Objective To quantify if risk selection in Dutch obstetric care (by midwives) results in a true low risk population in primary care at the end of pregnancy. This is an essential quality of care indicator as a distinction is made between primary care for low risk pregnancies by independently practicing community midwives, and secondary/tertiary care for high risk pregnancies by obstetricians.

Methods All singleton pregnancies (≥22 weeks' gestation, 2000-2007, n=1,407,387) from The Netherlands Perinatal Registry were selected. We defined high risk pregnancy as the presence of ≥1 Big4 morbidities, the main precursors of perinatal mortality: congenital anomalies, preterm birth, small for gestational age (SGA), or low Apgar score. Referral patterns of high risk pregnancies were studied during pregnancy and parturition; adequate risk selection implies no high risk pregnancies in primary care. Additionally, we applied a diagnostic test framework to study effectiveness of SGA selection (and referral) by defining true positives (referral of SGA), false positives (referral of non-SGA), false negatives (non-referral of SGA), and true negatives (non-referral of non-SGA). Sensitivity, specificity, negative predictive value (NPV), negative likelihood ratio (LR-), and false negative rate (FN) were determined for eight patient subgroups.

Results 59% of Big4 were referred during pregnancy, 19% during parturition; 22% remained in primary care. SGA 'test' characteristics differed considerably for subgroups (sensitivity 15%-59%, specificity 54%-87%, NPV 89%-97%, LR- 0.69-1.05, FN 3%-11%).

Conclusion Risk selection in Dutch obstetric care does not realise its aim of a true low risk group in primary care at the end of pregnancy. Methods for improvement are warranted.

INTRODUCTION

In The Netherlands perinatal mortality exceeds the European average.[1] The unique Dutch system of obstetric care has been regarded as a potential contributing factor.[2-4] This system is characterised by three risk-based levels of care. Primary care for low risk pregnancies is provided by independently practicing community midwives and a small percentage of general practitioners (GPs). Assumed low risk pregnant women can either opt for a home birth or a short-stay hospital birth under supervision of a community midwife. Secondary/ tertiary care for assumed high risk pregnancies is provided by obstetricians in hospitals.

Currently, approximately 80% of pregnant women start antenatal care in primary care.[5] Whenever risk factors (for adverse perinatal or maternal outcome) are present before pregnancy or arise during pregnancy or parturition, women shift from low risk to high risk and are referred to secondary care or from secondary to tertiary care, also during parturition. This ongoing risk assessment *during pregnancy* and *during parturition* is called 'risk selection'. In formal terms, the aim of risk selection is to identify and refer high risk pregnancies in order to obtain a true low risk group of pregnant women (expressed as high negative predictive value of risk selection) in the primary care setting.[5,6] Thus, risk selection adequacy is an essential quality of care indicator of the Dutch obstetric care system.

Although the effectiveness of risk selection in Dutch primary obstetric care has been studied, a nationwide systematic evaluation on the performance of the risk selection process is still absent.[3,4,6-11] The present nationwide retrospective study quantifies the performance of risk selection (*during pregnancy* and *during parturition*) by community midwives in terms of its ability to achieve a true low risk population at the end of pregnancy.

METHODS

Data

We selected data from all singleton pregnancies for the period 2000-2007 as registered in The Netherlands Perinatal Registry, which is subject to Dutch law regulations regarding confidentiality. In agreement with the World Health Organization (WHO) reporting guidelines, only pregnancies with a gestational age of ≥22 weeks were included.[12] The registry contains population-based information of >97% of all pregnancies in The Netherlands.[13] Source data are collected by 94% of midwives, 99% of gynaecologists and 68% of paediatricians (including 100% of Neonatal Intensive Care Unit paediatricians) as part

of their routine medical dossier; see website for detailed description www.perinatreg.nl.[13] The board of The Netherlands Perinatal Registry granted permission to use the anonymous registry data for this study. The Netherlands Perinatal Registry has been extensively described and used in several recent studies.[1,2,4,10,13-15]

Assignment of high risk

In Dutch obstetric care a pregnancy is considered a valid high risk, justifying referral, if an adverse perinatal outcome, adverse maternal outcome or combination of both is present or is to be expected. Indications for referral are listed in the so called 'List of Obstetric Indications' (in Dutch: *Verloskundige Indicatie Lijst*).[5] Community midwives are trained in the use of the 'List of Obstetric Indications' to detect (expected) high risks.

Judgment of adequacy of high and low risk assignment

As a retrospective measure to judge whether assumed low risk women truly were low risk, we used the prevalence of so-called Big4 morbidities as an indicator of risk status (gold standard).[2,15] From a detailed analysis of The Netherlands Perinatal Registry we know that four specific morbidities precede perinatal mortality in 85% of cases, the so-called Big4 morbidities.[2,15] These are: congenital anomalies (list defined), preterm birth (<37th week of gestation), small for gestational age (SGA, birthweight <10th percentile for gestational age[16]) or low Apgar score (<7, 5 minutes after birth). Congenital anomalies are registered postpartum through a standard coding system with eight different organ systems, and further distinction into 51 specific and 20 more global categories. By using remnant Big4 morbidity among assumed low risk women as yard stick we focus on undetected risks which are relevant to perinatal mortality. This focus by definition does not include any unexpected adverse maternal outcome in low risk women.

Effectiveness of risk selection

The primary outcome in quantifying the effectiveness of risk selection is the Big4 (high risk) prevalence at the end of pregnancy in primary care. In the theoretical perfect case Big4 morbidity is absent in assumed low risk pregnancies.

From this starting point, we utilise three methods to quantify effectiveness of risk selection:

Method 1. with a flow chart approach describing the proportional shift of women from primary care to secondary/tertiary care over the course of pregnancy, distinguishing

between two referral moments (*during pregnancy* and *during parturition*); risk selection should preferably take place *during pregnancy*;

Method 2. by comparing eight, mutually exclusive, patient subgroups in terms of the level of care at first antenatal booking (primary, secondary/tertiary care) and subsequent referral, where Big4 prevalence and perinatal mortality are observed in the various subgroups; preferably, perinatal mortality is only increased in the secondary/tertiary care group with referral of Big4 pregnancies *during parturition* being a rare event.

The eight subgroups were defined according to parity (primiparous/multiparous), ethnicity (Western/non-Western), and living in a deprived neighbourhood (yes/no, based on 4-digit zip codes and an official public list of 40 deprived zip code based neighbourhoods).[17] The eight groups presumably differ according to Big4 prevalence and care characteristics.

Method 3. by formal analysis of diagnostic performance of selection and referral of one Big4 category, i.e., SGA. The theoretical goal is to obtain a SGA-free population in the primary care setting, which is studied for the same eight patient subgroups as before. If the subgroup SGA prevalence matches the subgroup variation in test characteristics, then, the selected patient subgroup factors (parity/ethnicity/neighbourhood) are likely to be responsible for the between subgroup differences in test characteristics. However, if there is a discrepancy, other factors may explain the subgroup test characteristics variation factors, e.g., system related factors.

SGA was chosen as, together with congenital anomalies, it can be detected the easiest in the antenatal phase. Moreover, there is general consensus on (improving) detection of SGA because of the inherent increased risk for adverse outcome[18], and most congenital anomalies are now detected by routine ultrasound (introduced in 2007) at 20 weeks of gestational age.

Referral categories

To describe the referral process we defined five mutually exclusive categories:

I. First antenatal booking in secondary/tertiary care, no referral by definition, birth in hospital in secondary/tertiary care.

II. First antenatal booking in primary care, referral *during pregnancy*, birth in hospital in secondary/tertiary care;

III. First antenatal booking in primary care, referral *during parturition*, birth in hospital in secondary/tertiary care;

IV. First antenatal booking in primary care, no referral, home birth in primary care;

V. First antenatal booking in primary care, no referral, short-stay hospital birth in primary care.

Diagnostic performance of SGA referral and selection

For the third quantification of effectiveness of risk selection we applied a diagnostic test framework in which SGA selection (and referral) *during pregnancy* and *during parturition* are treated as a positive 'test result', whereas the presence of SGA at birth is regarded as the 'gold standard' outcome. The related 2x2 table is illustrated in table 2.1: 'A' represents true positives (referral of SGA), 'B' false positives (referral of non-SGA), 'C' false negatives (non-referral of SGA), and 'D' true negatives (non-referral of non-SGA).

The following five diagnostic test characteristics were determined applying method 3[19-21]:

- Sensitivity [A/(A+C)], the proportion of SGA cases referred;
- Specificity [D/(B+D)], the proportion of non-SGA cases which are not referred;
- Negative predictive value (NPV) [D/(C+D)], the proportion of non-referred women without a SGA baby;
- Negative likelihood ratio (LR-), [1-sensitivity/specificity], determines whether a negative 'test' result, i.e., no referral, decreases the probability of having a SGA baby for women who are not referred. A negative likelihood ratio ranging from 0 to <1 implies diagnostic value of the test, a value of 1 represents a test without diagnostic value ('similar to flipping a coin') regarding SGA selection and referral. The lower the LR-, the higher the diagnostic value, i.e., non-referral being associated with absence of SGA;
- SGA false negative rate (FN) [1-NPV], i.e., the proportion of non-referred women with a SGA baby (false negative).

Table 2.1 Main 2x2 table with actual referral to secondary/tertiary care treated as a positive 'test result' and the presence of SGA (small for gestational age) at birth treated as the 'gold standard' outcome

		SGA present at birth	
		Yes	No
Referral	Yes	A	B
	No	C	D

A: True positive / B: False positive / C: False negative / D: True negative.

The NPV, LR- and FN can be considered the most important test characteristics. These characteristics pertain to the least desirable situation, i.e., a woman with a high risk (e.g., SGA) pregnancy giving birth in primary care which is only intended for low risk pregnancies.

RESULTS

Process of risk selection

A total of 1,407,387 single pregnancies were analysed. Figure 2.1 displays the selection and referral of Big4 pregnancies by referral category in a flowchart; 15% of all pregnancies are Big4 pregnancies. The dark grey area represents the Big4 proportion in primary care (above the dashed line) and in secondary/tertiary care (below the dashed line). The dark grey area above the dashed line diminishes in width if Big4 pregnancies are referred from primary care to secondary/tertiary care *during pregnancy* or *during parturition*. Over the course of pregnancy, the proportion of Big4 pregnancies in primary care decreases from 14% at first antenatal booking to 6% in women giving birth at home and to 9% among women with a short-stay hospital delivery under the supervision of a community midwife.

Table 2.2 displays demographics and outcomes in the overall study population by the referral categories. Most women are multiparous (54%), between the ages of 20-35 years (84%), of Western origin (84%) and living in a non-deprived neighbourhood (94%). The largest referral group is the group of women referred during pregnancy (group II, n=466,415). In the group referred *during parturition* (group III), 70% of women are primiparous compared to 48% of women referred *during pregnancy* (group II). Group IV has the lowest risks, the group with the highest risk is group I.

Method 2: Big4 and perinatal mortality prevalence by patient subgroup and referral category

Table 2.3 shows the Big4 and perinatal mortality (overall 9.8 per 1,000) prevalence for the different subgroups in the five referral categories. Late Big4 referral, i.e., Big4 prevalence in women referred *during parturition* is not rare with a range of 14-19%; perinatal mortality, however, is relatively low ranging from 3.5 to 8.3 per 1,000 births.

Overall, there are large differences in perinatal mortality and Big4 prevalence between subgroups and referral categories: in all referral categories, primiparous women have

Figure 2.1 Referral and risk selection in Dutch obstetric care in absolute numbers; effectiveness of risk selection is illustrated as the proportion of Big4 pregnancies in dark grey. The different referral categories are light grey; from left to right referral categories I to V.

Table 2.2 Demographics and outcome in the study population by the five different referral categories

	Referral categories					
Referral category	I	II	III	IV	V	
First booking	Secondary/ Tertiary care	Primary care	Primary care	Primary care	Primary care	
Birth	Secondary/ Tertiary care	Secondary/ Tertiary care	Secondary/ Tertiary care	Primary care: home	Primary care: hospital	
Referral	No referral	During pregnancy	During parturition	No referral	No referral	Total
N	253,867	466,415	190,345	336,282	160,478	1,407,387
Parity*						
Primiparous (%)	121,066 (48)	226,133 (48)	134,102 (70)	104,488 (31)	63,567 (40)	649,356 (46)
Multiparous (%)	132,794 (52)	240,121 (52)	56,193 (30)	231,719 (69)	96,826 (60)	757,653 (54)
Maternal age*						
<20 years (%)	4,725 (2)**	7,539 (2)	4,333 (2)	2,820 (1)	4,637 (3)	24,054 (2)
20-35 years (%)	206,984 (82)**	382,195 (82)	166,938 (88)	290,442 (86)	136,418 (85)	1,182,977(84)
>35 years (%)	42,158 (17)**	76,681 (16)	19,074 (10)	43,020 (13)	19,423 (12)	200,356 (14)
Ethnicity						
Western (%)	214,332 (84)	387,985 (83)	157,245 (83)	308,425 (92)	112,775 (70)	1,180,762 (84)
Non-Western (%)	39,535 (16)	78,430 (17)	33,100 (17)	27,857 (8)	47,703 (30)	226,625 (16)

(*continued*)

Table 2.2 Continued

	Referral categories					
Referral category	I	II	III	IV	V	
First booking	Secondary/ Tertiary care	Primary care	Primary care	Primary care	Primary care	
Birth	Secondary/ Tertiary care	Secondary/ Tertiary care	Secondary/ Tertiary care	Primary care: home	Primary care: hospital	
Referral	No referral	During pregnancy	During parturition	No referral	No referral	Total
Neighborhood						
Non-deprived	237,836 (94)	437,358 (94)	177,652 (93)	324,477 (96)	143,699 (90)	1,321,022 (94)
Deprived	16,031 (6)	29,057 (6)	12,693 (7)	11,805 (4)	16,779 (10)	86,365 (6)
Congenital anomalies						
No (%)	245,554 (97)	447,192 (96)	185,515 (97)	332,647 (99)	158,329 (99)	1,369,237 (97)
Yes (%)	8,313 (3)	19,223 (4)	4,830 (3)	3,635 (1)	2,149 (1)	38,150 (3)
Preterm birth						
No (%)	223,946 (88)	426,324 (91)	177,404 (93)	334,920 (100)	157,945 (98)	1,320,539 (94)
Yes (%)	29,921 (12)	40,091 (9)	12,941 (7)	1,362 (0)	2,533 (2)	86,848 (6)
Small for gestational age						
No (%)	231,494 (91)	425,522 (91)	178,062 (94)	321,107 (95)	150,532 (94)	1,306,717 (93)
Yes (%)	22,373 (9)	40,893 (9)	12,283 (6)	15,175 (5)	9,946 (6)	100,670 (7)

* Totals do not always add up due to missing values (parity, 378 missings, maternal age, 284 missings).
** Percentages add up to, 101% because of rounding.

Table 2.3 Prevalence of Big4 morbidities and perinatal mortality by the five referral categories in different subgroups (based on parity, ethnicity and neighbourhood). Total number of pregnancies (N=1,407,009) does not include 378 parity missings

		Referral categories																	
Referral category		I			II			III			IV			V					
First booking		Secondary/Tertiary care			Primary care			Primary care			Primary care			Primary care					
Birth		Secondary/Tertiary care			Secondary/Tertiary care			Secondary/Tertiary care			Primary care: home			Primary care: hospital					
Referral		No referral			During pregnancy			During parturition			No referral			No referral			Total		
		N	Big4*	†**	N	Big4*	†**	N	Big4*	†**	N	Big4*	†**	N	Big4*	†**	N	Big4*	†**
Subgroups																			
Primiparous																			
Western	Non-deprived	101,876	25%	20.3	188,323	24%	13.8	110,669	16%	3.7	94,435	9%	0.8	44,882	12%	5.9	540,185	19%	10.0
	Deprived	3,284	28%	26.5	6,379	27%	16.3	3,719	18%	3.5	2,597	8%	0.8	2,063	15%	3.4	18,042	21%	11.8
Non-Western	Non-deprived	12,293	30%	34.2	24,570	27%	17.6	15,316	18%	5.3	5,947	12%	1.7	12,302	14%	5.1	70,428	22%	14.3
	Deprived	3,613	30%	39.6	6,861	29%	18.4	4,398	19%	5.5	1,509	13%	2.7	4,320	14%	2.8	20,701	23%	14.9
Multiparous																			
Western	Non-deprived	106,312	18%	19.6	188,639	15%	11.7	41,725	14%	6.2	208,150	5%	0.6	63,800	6%	3.6	608,626	11%	8.1
	Deprived	2,858	27%	33.9	4,522	20%	13.0	1,097	16%	7.3	3,179	6%	1.9	1,993	8%	2.0	13,649	16%	12.7
Non-Western	Non-deprived	17,348	23%	31.5	35,683	18%	15.6	9,897	15%	7.0	15,876	7%	0.6	22,640	7%	2.5	101,444	15%	12.2
	Deprived	6,276	25%	33.5	11,277	20%	17.0	3,474	15%	8.3	4,514	6%	0.9	8,393	8%	1.4	33,934	15%	13.2
Total		253,860	22%	22.3	466,254	20%	13.5	190,295	16%	4.7	336,207	6%	0.7	160,393	9%	4.0	1,407,009	15%	9.8

* Big4 morbidities in percentage of the total amount of pregnancies in that subgroup.
** Perinatal mortality (from 22 weeks of gestational age until 7 days postpartum) per 1,000 births.

higher Big4 prevalences compared to multiparous women (range 8-30% versus 5-27% for multiparous women). Perinatal mortality, however, is lower in primiparous women referred *during parturition* compared to multiparous women (3.5-5.5 vs. 6.2-8.3 per 1,000 births, respectively). In almost all subgroups and referral categories, non-Western women have higher Big4 and perinatal mortality prevalences; outcome differences between living in a non-deprived versus living in a deprived neighbourhood are smaller.

Method 3: SGA selection *during pregnancy*

Table 2.4 shows the SGA prevalence before and after selection (and subsequent referral), and diagnostic test characteristics for SGA selection and referral *during pregnancy* and *during parturition*. Of all women exposed to the selection and referral process *during pregnancy* (referral categories II to V), non-Western primiparous women in deprived neighbourhoods have the highest SGA prevalence (13%). For all subgroups, SGA prevalence is lower after selection compared to before. Sensitivity and specificity of SGA selection *during pregnancy* ranges from 49% to 59% and from 58% to 63% respectively. The range of NPV is 89-97%. The negative likelihood ratio (LR-) ranges from 0.69 to 0.85, indicating that that the SGA prevalence decreases (after the selection / 'test') in the group of women who are not referred *during pregnancy*. However, the values are close to 1 (which would imply a test without diagnostic value) which implies a modest discriminating value. The FN ranges from 3-11% depicting the percentage of non-referred women having SGA babies.

Method 3: SGA selection *during parturition*

Of all women exposed to the selection and referral process *during parturition* (referral categories III to V) primiparous non-Western women have the highest SGA prevalence (11%). For all subgroups, there is no difference in SGA prevalence before and after selection. Sensitivity and specificity of SGA selection *during parturition* range from 15% to 45% and from 54% to 87%, respectively. For primiparous women the sensitivity is higher than for multiparous women. Specificity, on the other hand, is higher for multiparous women than for primiparous women. The NPV is similar to that for the SGA selection process *during pregnancy* 89-97%. The LR- ranges from 0.97 to 1.05 implying no diagnostic value of risk selection for SGA *during parturition*. The FN ranges from 3-11%.

Table 2.4 Subgroup SGA prevalence before ('SGA% BF') and after ('SGA% AF') selection, and diagnostic test characteristics*; both for selection and referral of SGA pregnancies *during pregnancy* and *during parturition*

		Selection during pregnancy							Selection during parturition						
		SGA% BF	SGA% AF	SENS	SPEC	NPV	LR-	FN	SGA% BF	SGA% AF	SENS	SPEC	NPV	LR-	FN
Subgroups															
Primiparous															
Western	Non-deprived	9%	7%	53%	58%	93%	0.82	7%	7%	7%	42%	56%	93%	1.05	7%
	Deprived	11%	9%	53%	58%	91%	0.80	9%	9%	9%	44%	56%	91%	1.01	9%
Non-Western	Non-deprived	12%	11%	50%	59%	89%	0.85	11%	11%	11%	45%	54%	89%	1.01	11%
	Deprived	13%	11%	49%	61%	89%	0.83	11%	11%	11%	44%	57%	89%	0.99	11%
Multiparous															
Western	Non-deprived	4%	3%	52%	63%	97%	0.76	3%	3%	3%	15%	87%	97%	0.98	3%
	Deprived	7%	5%	59%	59%	95%	0.69	5%	5%	5%	18%	83%	95%	0.99	5%
Non-Western	Non-deprived	6%	5%	53%	58%	95%	0.81	5%	5%	5%	23%	80%	95%	0.97	5%
	Deprived	7%	5%	52%	60%	95%	0.79	5%	5%	5%	21%	79%	95%	1.00	5%
Total		7%	5%	52%	60%	95%	0.79	5%	5%	5%	33%	73%	95%	0.93	5%

* SENS: sensitivity, SPEC: specificity, NPV: negative predictive value, LR-: negative likelihood ratio, FN: SGA false negative rate.

DISCUSSION

Principal findings

To our knowledge, this is the largest study on the effectiveness of risk selection in Dutch primary obstetric care. The main focus was to examine whether risk selection realises its aim of a true low risk group of pregnant women by identifying and referring high risk pregnancies. Even though many Big4 pregnancies are referred, our results demonstrate that a true low risk population is never attained, with a Big4 prevalence of up to 15% in primary care (table 2.3), intended for low risk pregnancies only. Also, Big4 prevalence among late referrals (*during parturition*) was still substantial, ranging from 14% to 19%.

We further observed a suboptimal discrimination of SGA and non-SGA pregnancies (low sensitivity, LR- close to 1 in all subgroups) with a SGA prevalence of 3% to 11% still being born in primary care.

Moreover, we observed a discrepancy in the subgroup SGA prevalence and the subgroup variation in SGA selection test characteristics (table 2.4). This implies other factors to be responsible for the suboptimal test characteristics instead of the selected patient factors, e.g., system related factors. One may think of differences in availability of SGA screening methods.

Home birth versus short-stay hospital birth

In primary care, short-stay hospital births showed higher Big4 prevalence compared to home births (9% vs. 6%). This may reflect an unintentional selection process by either the midwife or self-selection by pregnant women, i.e., more healthy women appear to opt for home birth. This has also been observed in other studies.[22]

Preventability

From our results, the question arises whether the birth of a Big4 baby in a primary rather than secondary care setting is a preventable situation. As stated, congenital anomalies and SGA are better predictable than (spontaneous) preterm birth and a low Apgar score. In accepting that a NPV of 100% is not attainable, we actually state that Big4 deliveries in a primary care are to some extent inevitable in a system with different risk-based settings.

The subsequent question then refers to the observed 'setting safety' of a primary care setting. Both general consensus and the 'List of Obstetric Interventions' agree that a neonate with a Big4 morbidity is better off in a hospital setting under the care of an obstetrician/paediatrician.[23] The benefit of this so-called 'setting safety' may be related to availability of neonatological expertise, continuous fetal heart rate monitoring or advanced resuscitation equipment.[4,10]

As we showed less optimal SGA selection, and more referrals *during parturition* for primiparous women, we believe that all primiparous women should deliver in a hospital environment, either under supervision of a midwife (birth centre) or an obstetrician. This also follows from unequivocal evidence from previous studies[24-26], where others generally waive the primary care option.[24]

Strengths and limitations

The strengths of the current study include the nationwide approach and high coverage of pregnancies in The Netherlands Perinatal Registry over a long period of time. In addition, the use of Big4 morbidities, as the major precursors of perinatal mortality, allows for an easy-to-comprehend proxy measure of high risk pregnancy. [2,15] Also, the application of a diagnostic test framework allows the results to be interpreted objectively in a standardised way; comparisons with other diagnostic test studies can be easily made. Another strength pertains to the use of subgroups. It is interesting at the very least to see a discrepancy between the level of subgroup differences for SGA prevalence compared to the smaller subgroup differences in test characteristics.

This study also has limitations. Firstly, it is not possible to determine the exact indication for why women were referred because of the retrospective nature of The Netherlands Perinatal Registry. For our study objective, the effect of this limitation is limited as we focused on high risk births taking place in primary care, which is intended for low risk births. Our estimate of prevalence of high risk births in primary care is conservative as it is likely to be higher, providing it would have been possible to take into account all referral indications. Another possible limitation is that with the Big4 approach, maternal and other non-Big4 related risks are disregarded. This problem also appears to be limited as the majority of referral indications according to the 'List of Obstetric Interventions' pertain to fetal/neonatal risks alone.[5] Finally, the impact of routine ultrasound examination at 20 weeks of gestational age (introduced in 2007) on congenital abnormality rates and perinatal mortality rates due to second trimester abortions cannot be evaluated in our 2000-2007 dataset.

Previous studies on risk selection in the Dutch system

Our findings contradict most previous studies on the effectiveness of risk selection in Dutch primary obstetric care, stating that risk selection is effective.[6-9] In contrast with our study, these studies took into account maternal morbidity and obstetric interventions (e.g., caesarean section). However, these studies have been conducted some time ago, are restricted to smaller study groups, specific regions or did not evaluate the risk selection process systematically.[6-9] Another limitation of previous studies is that they defined the effectiveness of risk selection not only as prevention of adverse perinatal outcomes but also as prevention of obstetric interventions such as a caesarean section.[7-9] While we recognise that the assessment of risk selection must be weighed against the risk of possibly unnecessary obstetric interventions, the primary goal of adequate risk selection and subsequent referral is to prevent adverse perinatal and/or maternal outcomes, the prevention of (unnecessary) obstetric interventions being an important secondary goal.[6,9,27]

Several reports have expressed concern on the effectiveness of risk selection in Dutch primary obstetric care.[2,4,10] A recent Dutch study revealed that in 43% of Neonatal Intensive Care Unit (NICU) admissions the pregnancy had been indicated as low risk, and thus parturition had started in primary obstetric care.[10] Furthermore, infants of pregnant women at supposedly low risk whose labor started in primary care had a significantly higher delivery related perinatal mortality risk than the infants of assumed high risk women whose labor started in secondary care (relative risk 2.33, confidence interval 1.12-4.83).[4] Infants of women who were referred *during parturition* had a 3.66 times higher risk of delivery related perinatal mortality than infants of women who started labor in secondary care, and a 2.5-fold higher risk of NICU admission.[4] These studies emphasise that the level of healthcare provision could be improved for a proportion of supposedly low risk pregnant women at the onset of labor. Whether the delay in referral is related to late diagnosis (no continuous fetal heart rate monitoring *during parturition* in primary care), transport to hospital or assessment ('primary care is supposedly low risk'), is yet unclear.[4,10]

Possible implications

Our results demonstrate that the aim of risk selection in Dutch primary obstetric care is suboptimally attained. We propose some directions of improvement. As stated in the 'List of Obstetric Interventions', risk selection is currently exclusively done by primary care community midwives. Possible improvements could be the increase of midwives' competence and capabilities, or introduction of a checklist-based standardised risk

selection strategy at first antenatal booking. However, we believe that the required pace of change is more likely to be achieved through a combination of the latter with 'shared care': better cooperation between midwives and obstetricians who are jointly responsible for the determination of a woman's risk status, thereby joining their expertise which is either physiology-based (midwives) or pathology-based (obstetricians).[28] Shared obstetric care has already been implemented in some form in other Western countries such as Australia and the United Kingdom.[29-31] One study demonstrated a 27% increase in the detection rate of intrauterine growth restriction for women receiving shared obstetric care as opposed to conventional obstetric care.[32] For more generalisable results however, a study to evaluate different shared care strategies has to be conducted in The Netherlands because of the unique system of obstetric care. We believe that our recommendation for shared care also applies to countries which are considering or already have an obstetric care system with features similar to the Dutch system, such as Canada.[33-36]

REFERENCES

1. Mohangoo AD, Buitendijk SE, Hukkelhoven CW, et al. Higher perinatal mortality in The Netherlands than in other European countries: the Peristat-II study. Ned Tijdschr Geneeskd. 2008;152:2718-2727.
2. Bonsel GJ, Birnie E, Denktas S, Poeran J, Steegers EAP. Dutch report: Lijnen in de Perinatale Sterfte, Signalementstudie Zwangerschap en Geboorte 2010. Rotterdam: Erasmus MC; 2010.
3. Sheldon T. Obstetric care must change if Netherlands is to regain reputation for safe childbirth. BMJ. 2008;336:c239.
4. Evers AC, Brouwers HA, Hukkelhoven CW, et al. Perinatal mortality and severe morbidity in low and high risk term pregnancies in the Netherlands: prospective cohort study. BMJ. 2010;341:c5639.
5. Amelink-Verburg MP, Buitendijk SE. Pregnancy and labour in the Dutch maternity care system: what is normal? The role division between midwives and obstetricians. J Midwifery Womens Health. 2010;55:216-225.
6. Bais JMJ, Pel M. The basis of the Dutch obstetric system: risk selection. Eur Clinics Obstet Gynaecol. 2007;2:209-212.
7. Eskes M. Het Wormerveeronderzoek. Meerjarenonderzoek naar de kwaliteit van de verloskundige zorg rond een vroedvrouwenpraktijk. Thesis 1989, University of Amsterdam.
8. Berghs G, Spanjaards E. De normale zwangerschap: bevalling en beleid. Thesis 1988, Catholic University Nijmegen.
9. Bais JMJ. Risk selection and detection. A critical appraisal of the Dutch obstetric system. Thesis 2004, University of Amsterdam.
10. Evers AC, van Leeuwen J, Kwee A, et al. Mortality and morbidity among full-term neonates in a neonatal intensive care unit in the Utrecht region, the Netherlands. Ned Tijdschr Geneeskd. 2010;154:118.
11. Sheldon T. Prompt access to expert care is among advice to reduce perinatal deaths in Netherlands. BMJ. 2010;340:c277.
12. World Health Organization. Neonatal and perinatal mortality. Country, regional and global estimates. Geneva: WHO, 2006. (online: http://whqlibdoc.who.int/publications/2006/9241563206_eng.pdf).
13. Perinatal care in The Netherlands 2007 (in Dutch: Perinatale zorg in Nederland 2007). Utrecht: The Netherlands Perinatal Registry; 2009.
14. Tromp M, Meray N, Ravelli AC, Reitsma JB, Bonsel GJ. Medical Record Linkage of Anonymous

Registries without Validated Sample Linkage of the Dutch Perinatal Registries. Stud Health Technol Inform. 2005;116:125-130.

15. Van der Kooy J, Poeran J, de Graaf JP, et al. Planned home compared with planned hospital births in the Netherlands: intrapartum and early neonatal death in low-risk pregnancies. Obstet Gynecol. 2011;118:1037-1046.

16. Kloosterman GJ. On intrauterine growth. Int J Gynaecol Obstet. 1970;8:895-912.

17. Gent, van WPC. Realistic regeneration -housing contexts and social outcomes of neighbourhood interventions in Western European cities-; chapter 4. Thesis 2009, University of Amsterdam. (online: http://dare.uva.nl/document/153004).

18. Figueras F, Gardosi J. Intrauterine growth restriction: new concepts in antenatal surveillance, diagnosis, and management. Am J Obstet Gynecol. Apr 2011;204:288-300.

19. Altman DG, Bland JM. Diagnostic tests 2: Predictive values. BMJ. 1994;309:102.

20. Altman DG, Bland JM. Diagnostic tests. 1: Sensitivity and specificity. BMJ. 1994;308:1552.

21. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. BMJ. 2004;329:168-169.

22. Kennare RM, Keirse MJ, Tucker GR, Chan AC. Planned home and hospital births in South Australia, 1991-2006: differences in outcomes. Med J Aust. 2010;192:76-80.

23. National Institute for Health and Clinical Excellence (see Antenatal Care: Routine Care for the Healthy Pregnant Woman). (online: http://www.nice.org.uk/nicemedia/live/11947/40115/40115.pdf).

24. Chervenak FA, McCullough LB, Brent RL, Levene MI, Arabin B. Planned home birth: the professional responsibility response. Am J Obstet Gynecol. 2013;208:31-38.

25. Brocklehurst P, Hardy P, Hollowell J, et al. Perinatal and maternal outcomes by planned place of birth for healthy women with low risk pregnancies: the Birthplace in England national prospective cohort study. BMJ. 2011;343:d7400.

26. De Neef T, Hukkelhoven CW, Franx A, van Everhardt E. Uit de lijn der verwachting. Nederl Tijdschrift Obstet Gynaecol 2009;122:34-342.

27. National Institute for Health and Clinical Excellence. Intrapartum care: care of healthy women and their babies during childbirth. CG55. 2007. (online: www.nice.org.uk/nicemedia/pdf/IPC-NICEGuidance.pdf).

28. Posthumus AG, Scholmerich VL, Waelput AJ, et al. Bridging Between Professionals in Perinatal Care: Towards Shared Care in The Netherlands. Matern Child Health J. 2012 (in press).

29. Halliday J. WUDWAW 'Who Usually Delivers Whom and Where', Perinatal Data Collection Unit. Victoria, Australia: Victorian Government Department of Human Services; 1999.

30. Hampson JP, Roberts RI, Morgan DA. Shared care: a review of the literature. Fam Pract. 1996;13: 264-279.

31. Tucker J, Florey CD, Howie P, McIlwaine G, Hall M. Is antenatal care apportioned according to obstetric risk? The Scottish antenatal care study. J Public Health Med. 1994;16:60-70.

32. Chan FY, Pun TC, Tse LY, Lai P, Ma HK. Shared antenatal care between family health services and hospital (consultant) services for low risk women. Asia Oceania J Obstet Gynaecol. 1993; 19:291-298.

33. De Vries R, Benoit C, Van Teijlingen E, Wrede S. Birth by design: pregnancy, maternity care and midwifery in North America and Northern Europe. New York: Routledge, 2001.

34. Hatem M, Sandall J, Devane D, Soltani H, Gates S. Midwife-led versus other models of care for childbearing women. Cochrane Database Syst Rev. 2008:CD004667.

35. Bick D. Enhancing safety in the maternity services: a greater role for midwife-led care? Midwifery. 2009;25:1-2.

36. Royal College of Obstetricians and Gynaecologists/ Royal College of Midwives. Home births. Joint statement No.2; 2007 (online: http://www.rcog.org.uk/files/rcog-corp/uploaded-files/JointStatmentHomeBirths2007.pdf).

3

Planned home compared with planned hospital births in The Netherlands: intrapartum and early neonatal death in low-risk pregnancies

J. van der Kooy, J. Poeran, J.P. de Graaf, E. Birnie,
S. Denktaş, E.A.P. Steegers, G.J. Bonsel

Obstetrics and Gynecology. 2011;118:1037-46.

ABSTRACT

Objective The purpose of our study was to compare the intrapartum and early neonatal mortality rate of planned home birth vs. planned hospital birth in community midwife-led deliveries, after case mix adjustment.

Methods Perinatal outcome of 679,952 low-risk women was obtained from the Dutch Perinatal Registry (2000-2007). This group represents all women who had a choice between home and hospital birth. Two different analyses were performed; natural prospective approach (intention-to-treat like analysis) and perfect guideline approach (per-protocol like analysis). Unadjusted and adjusted odds ratios were calculated. Case mix was based on the presence of at least one of the following: congenital abnormalities, small for gestational age, preterm birth, or low Apgar score. We also investigated the potential risk role of intended place of birth. The technique used was multivariable stepwise logistic regression.

Results Intrapartum and neonatal death 0-7 days was observed in 0.15% of planned home vs. 0.18% in planned hospital births (crude RR 0.80 95%CI 0.71-0.91). After case mix adjustment, the relation is reversed, showing non-significant increased mortality risk of home birth (OR 1.05 95%CI 0.91-1.21). In certain subgroups additional mortality may arise at home if risk conditions emerge at birth (up to 20% increase).

Conclusion Home birth, under routine conditions, is generally not associated with increased intrapartum and early neonatal death, yet in subgroups additional risk cannot be excluded.

INTRODUCTION

The debate on the safety of home births continues in the literature as recently addressed in the Lancet.[1] In the Netherlands, approximately 50% of women give birth under the supervision of a community midwife. The community midwives are independent health care professionals in the Netherlands, operating either solely or in group practices. The proportion of home birth deliveries in the Netherlands has steadily decreased over the last decade but is currently stable at 25% of all births. Several Anglo-Saxon countries are considering the reintroduction of home births, based on recent claims of sufficient safety.[2] The reverse trend is observed in the Netherlands, where the debate has intensified since the national perinatal mortality rate showed to be one of the highest in Europe.[3]

In the Dutch system, independently operating community midwives provide care for low- and medium-risk pregnant women (primary healthcare). High-risk pregnant women are referred to the gynaecologist for remaining ante- and intrapartum care. If no or only a few risk factors are present, women can stay with the midwife and decide where the delivery will take place: at home or in the hospital, both supervised by the community midwife. For pregnant women with so called 'medium-risk' delivery in hospital is obligatory but can still be under the supervision of the community midwife. A strict definition of medium risk, created and agreed upon by midwives and gynaecologists together, is defined in the Dutch guidelines.[4] The claimed benefits of planned home births include the reduction of maternal-fetal morbidity, a lower risk for unjustified medical interventions, and psychosocial advantages for the mother. These benefits may be counterbalanced by the disadvantages associated with a high intrapartum referral rate and an increased perinatal mortality, morbidity and long term negative effects.[5-11]

This paper re-addresses the Dutch evidence focusing on two critical features of previous analyses. First, previous studies compared outcomes after exclusion of pregnant women who in view of the delivery guidelines should have been referred to a gynaecologist. Second, previous studies did not apply case mix analysis, assuming risk equivalence of home and hospital groups.[5,9,12-18] Case mix may, however, differ across planned place of delivery, due to self selection or due to the midwife's proposal, with the healthiest and most affluent women receiving home birth (confounding the comparison by indication bias).[5 6,7,11,19-21]

The purpose of our study was to compare the intrapartum and early neonatal mortality rate of planned home birth vs. planned hospital birth in community midwife-led deliveries, after case mix adjustment. We compared a natural prospective approach without ex post exclusion of unsuitable midwife cases (intention-to-treat like), with the conventional

approach based on a theoretical midwife population under perfect guideline adherence (per-protocol like). We hypothesised that while in general no difference may exist between home and hospital outcomes, for specific risk groups the hospital setting is protective as obstetrical and neonatal expertise and clinical facilities are directly available (so-called "setting safety").

METHODS

Data

The Netherlands Perinatal Registry (PRN) contains population-based information of 96% of all pregnancies in The Netherlands. Source data are collected by 95% of midwives, 99% of gynaecologists and 68% of paediatricians (including 100% of Neonatal Intensive Care Unit paediatricians).[3,22] (See website for detailed description: www.perinatreg.nl). We selected the records of all singleton pregnant women, under supervision of a community midwife at the onset of labour between 2000-2007 (693,592 women). The onset of labour was defined as spontaneous contractions or the spontaneous rupture of membranes by the PRN. Two subsets of pregnant women were further excluded from the original set of 693,592 women. First, 13,384 women with so called 'medium risk', for example women with a history of postpartum haemorrhage or obesity (BMI>30). Dutch guidelines prescribe a hospital delivery for these women which may be supervised by the community midwife. Secondly we excluded records were the data was incomplete (n=256).

The remaining women (n=679,952) were categorised according to intended place of birth, which usually is concordant with the observed place of birth either home or hospital. For some women the place was not decided until the onset of labour. This could be due to indifference on the part of the woman; or delayed antepartum care. The intended place was then coded 'unknown'. This yielded 3 intention groups: home, hospital, and, unknown.

Outcome measures, maternal and neonatal risk factors

Outcome was defined as intrapartum and early neonatal mortality , i.e. (I) intrapartum death, (II) neonatal death up to 24 hrs, and (III) neonatal death up from 1 day to 7 days post partum. In our low risk group under midwife supervision, mortality beyond 8 days is rare, and regarded to be unrelated to place of delivery. The PRN does not include long term child outcomes for example psychomotor development and behavioural function.

Maternal risk factors were parity (nulliparous vs. multiparous), age, ethnicity (Western/non-Western; based on a more refined classification in the registry), and living in a deprived neighbourhood (yes/no, based on 4-digit zip-codes and a public list of deprived, zip-code based, neighbourhoods issued by the Dutch government).

Detailed risk information is unavailable in national registries. Case mix of any defined group of women was primarily represented by the prevalence (single or combined) of Big4 conditions (see below). From detailed analysis of the complete perinatal dataset of the same Netherlands Perinatal Registry (PRN), years 2000-2007, (1.25 million records)[23], it appeared that the presence of any of 4 conditions preceded perinatal mortality in 85% of cases. These conditions were defined as; congenital abnormalities (list defined), small for gestational age (SGA, birth weight below the 10th percentile for gestational age, gender and parity specific), preterm birth (< 37th week of gestation) or low Apgar score (< 7, measured 5 minutes after birth). We will continue to refer to these 4 conditions as the Big4. The main results of this detailed analysis are found in figure 3.1.

In our current analysis these so called Big4 represent an objective estimate of the risk challenge at birth. The preventability of their occurrence, either antenatally or during delivery, is not at issue. Here we intentionally use it as a risk indicator, an explanatory factor at onset. By doing so, we ignore differential management effects of setting on the emergence of these Big4, in particular low Apgar, should they exist.

Data analysis

As primary analysis we present the results of the natural prospective approach (NPA), resembling an intention-to-treat analysis. For comparison we added a perfect guideline approach (PGA), resembling a per-protocol analysis. The NPA approach establishes, within observational constraints, the intrapartum and early neonatal death of planned home versus planned hospital births. It stems from the viewpoint of a pregnant woman starting birth under supervision of a midwife (the denominator is n=679,952). The natural approach thus includes spontaneous preterm labour since to some extent this group was not referred to the gynaecologist during labour or was referred late during (home) delivery. Therefore a direct setting effect (admission to hospital at the onset of labour) may be visible to the advantage of the hospital. Furthermore indirect setting effects may be present, for example the timing of referral.

PGA includes the subset of women within the NPA population, who *in retrospect* were compliant with the guidelines which define low risk at the onset of labour and therefore allowed to choose between a home or hospital birth under supervision of a midwife.

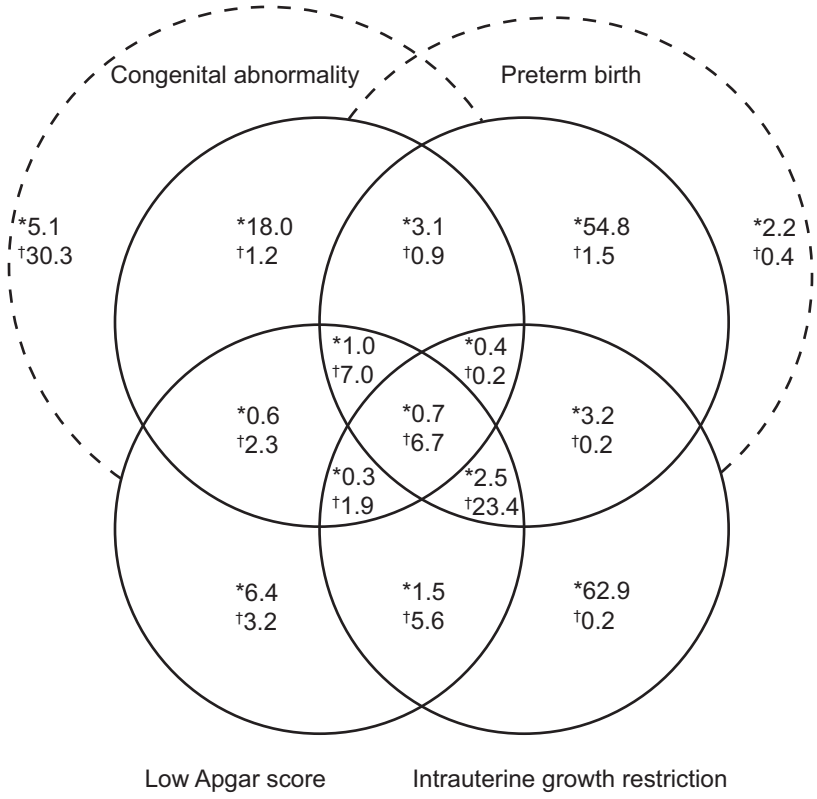


Figure 3.1 Perinatogram illustrating in a Venn diagram the relationship between (combinations of) Big 4 morbidities and perinatal mortality defined as death from 22 weeks of gestational age until 7 days postpartum. In 85% of all cases of perinatal mortality, one or more Big4 morbidities are present; for instance, a low Apgar score combined with preterm birth occurs in 30.3% of all cases of perinatal mortality. *Prevalence per 1,000 births of separate and combined Big 4 morbidities and their contribution to all cases of perinatal mortality (†percentage); this adds up to 85% of all cases of perinatal mortality. The dashed circles connect low Apgar score with preterm birth and congenital abnormality with intrauterine growth restriction.

Non-compliance exists if a high risk condition was already detectable at the onset of labour. These conditions applied to women with a gestational age <37 or >41 weeks, prolonged rupture of membranes (>24hr) and intrauterine death with unclear timing relative to onset of labour (see figure 3.2). PGA (n=602,331) still included undetected SGA and congenital malformations that emerge at birth, as detection failure cannot be regarded as non-compliance from the viewpoint of current guidelines.

First we compared characteristics of the NPA and PGA populations by intended place of birth (t-tests for comparisons). Then we investigated the potential risk role of intended place

of birth by a set of predefined nested multivariable logistic regression models (stepwise analysis; inclusion $p<0.05$; exclusion $p>0.10$) where we added maternal and neonatal (case mix) explanatory variables. For these variables, hospital birth was set as the reference. All stepwise analyses were repeated with a forward and backward approach, and finally forced inclusion of predictive variables ($p<0.05$). Risk factor coefficients were only shown

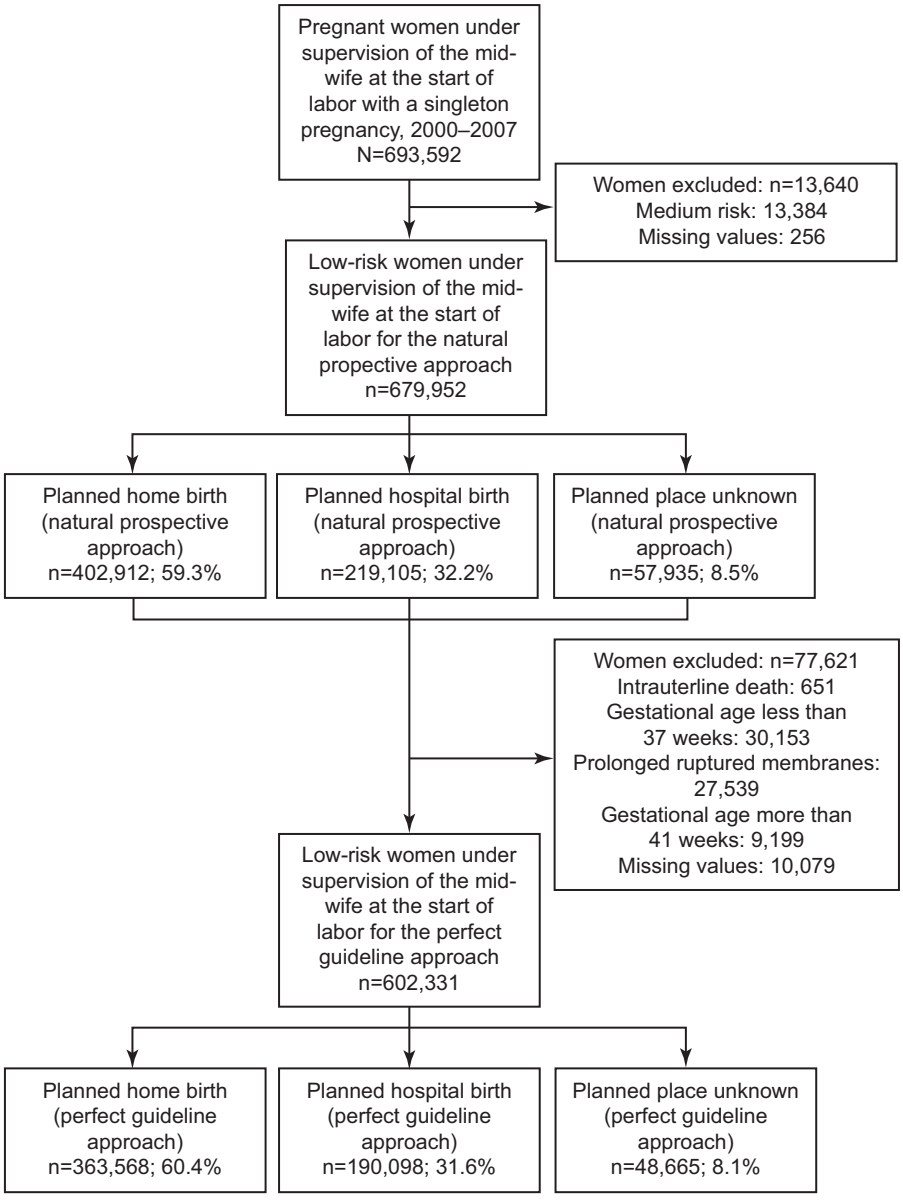


Figure 3.2 Flow of women through the study.

in case of significance $p<0.05$. Results across the three approaches were similar unless stated otherwise.

We graphically described the crude mortality of the planned home and planned hospital population, for the series of populations which result from successive exclusion of women meeting a criterion for non-compliance (figure 3.3; dotted lines). This successive exclusion through non-compliance criteria gradually transforms the NPA population into the PGA population. If the mortality rate of a non-compliance group is average, home and hospital mortality rates do not change on its exclusion. If the rates decrease at a different gradient (e.g. hospital steeper than home, as after exclusion of pregnancy duration < 36 weeks) this may point to either differential prevalence of the non-compliance factor (as here), or to differential case fatality by setting where the largest mortality decrease is observed in the setting with the highest case fatality (interpretable as lowest setting safety).

To support this interpretation, we first divided the crude mortality of the home and hospital group by the respective prevalence of Big4 conditions to obtain case mix adjustment. This assumes Big4 prevalence to be a suitable risk indicator at the group level. Subsequent division of the resulting home/Big4 mortality ratio by the hospital/Big4 mortality yields an index (Big4 adjusted homebirth mortality index; figure 3.3; black line). If this index is 100%, then relative mortality in home births and hospital births is equal. If the index is for example 120%, then home births have 20% excess mortality taking our case mix differences into account. Combining crude mortality changes with index changes allows for tentative interpretation of setting effects.

RESULTS

Table 3.1 describes the baseline characteristics of both the NPA and PGA populations (n=679,952 vs. 602,331).

In both the NPA and PGA populations about 60% of women planned a home delivery and about 32% planned a hospital delivery. Compared to women who planned birth in the hospital or with unknown location, the women with planned home birth were more likely to be multiparous, 25 years or older, of Dutch origin and to live in a privileged neighbourhood (all of which are favourable conditions). In home birth women, neonatal case mix compared also favourably. Premature delivery was less common, as was the prevalence of a Big4 condition (NPA home birth 8.7% vs. hospital 10.8% vs. unknown 10.5%; PGA home birth 6.5% vs. hospital 8.2% vs. unknown 7.5%; $p<0.001$ in both cases).

Intrapartum and early neonatal mortality was 1099/679,952=1.62 ‰ in the NPA women and 551/602,331=0.91‰ in PGA women. Mortality was lower in women who were multiparous, between 24-34 year, of Dutch origin, or living in a privileged neighbourhood (both NPA and PGA), see table 3.1. Within the group with intrapartum and early neonatal mortality, Big4 conditions were found in 792 of the 1099 deaths (72.1%) in the NPA women, compared to 290 out of 551 deaths (52.6%) in the PGA group.

In the NPA population, crude mortality risk was significantly lower for women who planned to give birth at home (RR 0.80 95%CI 0.71-0.91) and for women with unknown intention (RR 0.96 95%CI 0.77-1.19) compared to those who intended to give birth in hospital ($P<0.05$) (see table 3.2). All maternal and neonatal risk factors, except living in a deprived neighbourhood, showed significant effect sizes in agreement with the expected direction. Mortality was significantly increased in infants with a Big4 outcome, especially in infants with multiple Big4 conditions (RR 168.9 95%CI 148.9-191.4).

The nested multivariable logistic regression analysis showed that in the presence of adjusting maternal factors only (model 2), the intended place of birth had no significant impact on outcome. The maternal factors showed risks similar to the univariable (crude) analysis. The addition of Big4 case mix adjustment (model 3) showed the intended place of birth to be a significant co-variable, yet the contrast of planned home birth (OR 1.05 95%CI 0.91-1.21) vs. hospital birth (reference=1) turned out to be non-significant. The effect of maternal risk factors was affected to a limited degree by the introduction of Big4 case mix.

We repeated the analysis for the PGA population (table 3.3). The results of the crude analysis were close to the NPA analysis. However, the effect of ethnic background was considerably stronger in the PGA population. In all analyses the intended place of birth showed a consistent significant impact on intrapartum and early neonatal mortality, yet the contrast between home and hospital birth never reached statistical difference. After Big4 case mix adjustment home birth showed a non-significant increased risk (OR 1.11 95%CI 0.93-1.34).

Figure 3.3 describes the crude mortality risk (left Y-axis) and the Big4 adjusted home birth mortality index (right Y-axis), where each dot represents the mortality risk results after the group listed on the X-axis has been excluded from the population.

The crude mortality (dotted lines) initially shows a difference in favour of home delivery (home: 0.18% vs. hospital: 0.22%), which converges towards a much lower average level if premature births are excluded. Further exclusions lower the crude mortality rate, leaving the small difference almost unaffected. The mortality index (black line) shows a distinct change from an initial level of about 100% towards about 120% after exclusion of the

Table 3.1 Characteristics and outcomes of women in primary care at the onset of labour (natural prospective approach and perfect guideline approach)*

	Planned home birth		Planned hospital birth		Planned place unknown		Intrapartum and early neonatal death	
Variable	NPA	PGA	NPA	PGA	NPA	PGA	NPA	PGA
Parity†	402,912 (59.3)	363,568 (60.4)	219,105 (32.2)	190,098 (31.6)	57,935 (8.5)	48,665 (8.1)	679,952	602,331
Primiparous	171,986 (42.69)	148,082 (40.73)	104,249 (47.58)	88,110 (46.35)	26,254 (45.32)	21,047 (43.25)	614 (0.20)	283 (0.11)
Multiparous	230,926 (57.31)	215,486 (59.27)	114,856 (52.42)	101,988 (53.65)	31,681 (54.68)	27,618 (56.75)	485 (0.13)	268 (0.08)
Maternal age (y)†								
Younger than 19	4,036 (1.00)	3,502 (0.96)	6,713 (3.06)	5,770 (3.04)	1,190 (2.05)	910 (1.87)	42 (0.35)	13 (0.13)
20-25	34,661 (8.60)	30,787 (8.47)	32,617 (14.89)	28,669 (15.08)	6,823 (11.78)	5,611 (11.53)	133 (0.18)	65 (0.10)
25-34	296,128 (73.50)	267,408 (73.55)	142,597 (65.08)	124,071 (65.27)	39,526 (68.22)	33,583 (69.01)	693 (0.14)	348 (0.08)
Older than 35	68,087 (16.90)	61,871 (17.02)	37,178 (16.97)	31,588 (16.62)	10,396 (17.94)	8,559 (17.59)	231 (0.20)	125 (0.12)
Ethnic background†								
Western	364,796 (90.54)	329,677 (90.68)	143,677 (65.57)	124,144 (65.31)	45,205 (78.03)	38,508 (68.80)	880 (0.16)	452 (0.09)
Non Western	38,116 (9.46)	33,891 (9.32)	75,428 (34.43)	65,954 (34.69)	12,730 (21.97)	17,461 (31.20)	219 (0.17)	99 (0.08)
Neighbourhood†								
Privileged neighbourhood	388,089 (96.32)	350,346 (96.36)	196,659 (89.76)	170,366 (89.62)	53,823 (92.90)	45,425 (93.34)	1,031 (0.16)	518 (0.09)
Underprivileged neighbourhood	14,823 (3.68)	13,222 (3.64)	22,446 (10.24)	19,732 (10.38)	4,112 (7.10)	3,240 (6.66)	68 (0.16)	33 (0.09)

Gestational age (wk)†								
Less than 34	2,409 (0.60)		1,702 (0.78)		583 (1.01)		370 (7.88)	
35-36	6,510 (1.62)		4,064 (1.85)		1,206 (2.08)		65 (0.55)	
37	15,203 (3.77)	13,622 (3.75)	9,603 (4.38)	8,468 (4.45)	2,497 (4.31)	2,187 (4.49)	56 (0.21)	51 (0.21)
38-41	368,926 (91.56)	349,946 (96.25)	193,816 (88.46)	181,630 (95.55)	49,585 (85.59)	46,478 (95.51)	548 (0.09)	500 (0.09)
More than 41	9,864 (2.45)		9,920 (4.53)		4,064 (7.01)		60 (0.25)	
Big4†								
Small for gestational age	18,786 (4.66)	17,089 (4.70)	13,114 (5.99)	11,604 (6.10)	3,081 (5.32)	2,665 (5.48)	71 (0.20)	59 (0.19)
Premature	8,090 (2.01)		5,117 (2.34)		1,547 (2.67)		92 (0.62)	0.00
Low Apgar score	1,692 (0.42)	1,483 (0.41)	1,180 (0.54)	959 (0.50)	289 (0.50)	228 (0.47)	97 (3.07)	86 (3.22)
Congenital abnomality	4,874 (1.21)	4,366 (1.20)	2,941 (1.34)	2,531 (1.33)	778 (1.34)	655 (1.35)	74 (0.86)	60 (0.79)
Combination Big4	1,648 (0.41)	693 (0.19)	1,279 (0.58)	453 (0.24)	391 (0.67)	92 (0.19)	458 (13.80)	85 (6.87)
Total Big4	35,090 (8.71)	23,631 (6.50)	23,631 (10.79)	15,547 (8.18)	6,086 (10.50)	3,640 (7.48)	792 (1.22)	290 (0.68)

NPA, natural prospective approach; PGA, perfect guideline approach.
Data are n (%).
* Totals may not add up to 100 because of rounding error.
† P<.001.

Table 3.2 Intrapartum and neonatal death 0-7 days in women who are in primary care at the onset of labour (natural prospective approach)

				Model 1			Model 2			Model 3		
	Total	Mortality	p	Crude RR	95% CI	p	Adjusted OR	95% CI	p	Adjusted OR	95% CI	p
Intended place of birth			<.05			<.05						<.05
Home	402,912	594 (0.15)		0.80	0.71–0.91		nie			1.05	0.91–1.21	
Hospital (ref)	219,105	403 (0.18)		1						1		
Unknown	57,935	102 (0.18)		0.96	0.77–1.19		nie			0.77	0.61–0.97	
Parity			<.001			<.001			<.001			
Primiparous	302,489	614 (0.20)		1.58	1.40–1.78		1.67	1.47–1.89		nie		
Multiparous (ref)	377,463	485 (0.13)		1			1					
Maternal age (y)			<.001			<.001			<.001			<.001
Younger than 19	11,939	42 (0.35)		2.43	1.78–3.31		1.80	1.31–2.48		1.67	1.17–2.38	
20-25	74,101	133 (0.18)		1.24	1.03–1.49		1.03	0.85–1.24		0.92	0.75–1.13	
25-34 (ref)	478,017	693 (0.14)		1			1			1		
Older than 35	115,661	231 (0.20)		1.38	1.19–1.60		1.56	1.34–1.81		1.44	1.23–1.68	
Ethnic background			<.001			<.001			<.05			<.05
Western (ref)	571,185	880 (0.15)		1			1			1		
Non-Western	108,767	219 (0.20)		1.31	1.13–1.52		1.32	1.14–1.54		1.21	1.02–1.45	

Neighbourhood						<.05				
Privileged neighbourhood (ref)	638,571	1,031 (0.16)		1						
Underprivileged neighbourhood	41,381	68 (0.16)		1.02	0.80–1.30		nie		nie	
Gestational age (wk)			<.001			<.001				<.001
Less than 34	4,694	370 (7.88)		88.08	77.44–100.18				15.74	12.56–19.72
35-36	11,780	65 (0.55)		6.17	4.77–7.97				2.11	1.56–2.86
37	27,303	56 (0.21)		2.29	1.74–3.02				2.26	1.71–2.99
38-41 (ref)	612,327	548 (0.09)		1					1	
More than 41	23,848	60 (0.25)		2.81	2.15–3.67				2.10	1.60–2.77
Big4			<.001			<.001				<.001
SGA	34,910	71 (0.20)		1.27	1.00–1.62				3.64	2.82–4.69
Prematurity	14,754	92 (0.62)		12.49	9.90–15.76					
Low apgar	3,161	97 (3.07)		21.35	17.28–26.38				53.41	42.55–67.04
Congenital abnomality	8,593	74 (0.86)		5.68	4.48–7.20				14.53	11.30–18.67
Combination Big4	3,318	458 (13.80)		168.85	148.94–191.41				68.17	55.27–84.08
No Big4	615,145	307 (0.05)		1						

RR, relative risk; CI, confidence interval; OR, odds ratio; nie, not in equation.
Data are n or n (%) unless otherwise specified.
Model 1: crude RR. Model 2: adjusted for maternal factors including intended place of birth, parity, age, ethnic background, and neighborhood. Model 3: adjusted for maternal factors and child factors; model 2 + gestational age and presence of Big 4.

Table 3.3 Intrapartum and neonatal death 0-7 days in women who are in primary care at the onset of labour (perfect guideline approach)

			Model 1			Model 2			Model 3		
	Total (n)	Mortality	Crude RR	95% CI	p	Adjusted OR	95% CI	p	Adjusted OR	95% CI	p
Intended place of birth					<.05			<.05			<.05
Home	363,568	344 (0.09)	0.99	0.83–1.18		1.02	0.85–1.23		1.11	0.93–1.34	
Hospital (ref)	190,098	182 (0.10)	1			1			1		
Unknown	48,665	25 (0.05)	0.54	0.35–0.81		0.54	0.36–0.83		0.57	0.37–0.86	
Parity					<.001			<.001			
Primiparous	257,239	283 (0.11)	1.42	1.20–1.67		1.52	1.28–1.82		nie		
Multiparous (ref)	345,092	268 (0.08)	1			1					
Maternal age (y)					<.001			<.001			<.05
Younger than 19	10,182	13 (0.13)	1.56	0.90–2.71		1.31	0.75–2.30		1.29	0.73–2.26	
20-25	65,067	65 (0.10)	1.22	0.94–1.59		1.11	0.84–1.45		1.08	0.83–1.41	
25-34 (ref)	424,915	348 (0.08)	1			1			1		
Older than 35	102,018	125 (0.12)	1.50	1.22–1.84		1.66	1.34–2.04		1.50	1.22–1.85	
Ethnic background											
Western (ref)	507,063	452 (0.09)	1								
Non-Western	94,717	99 (0.10)	1.17	0.94–1.46		nie			nie		

Neighbourhood									
Privileged neighbourhood (ref)	566,137	518 (0.09)	1						
Underprivileged neighbourhood	36,194	33 (0.09)	1.00	0.70–1.42		nie		nie	
Gestational age (wk)					<.001				<.001
37	24,277	51 (0.21)	2.43	1.82–3.24			2.51	1.87–3.37	
38-41 (ref)	578,054	500 (0.09)	1				1		
Big3					<.001				<.001
SGA	31,358	59 (0.19)	4.03	3.04–5.35			4.28	3.22–5.68	
Low apgar	2,670	86 (3.22)	69.05	54.28–87.84			71.52	55.87–91.56	
Congenital abnomality	7,552	60 (0.79)	17.03	12.88–22.53			16.76	12.64–22.23	
Combination Big3	1,238	85 (6.87)	147.19	115.97–186.80			159.49	123.86–205.38	
No Big3	559,513	261 (0.05)	1				1		

RR, relative risk; CI, confidence interval; OR, odds ratio; nie, not in equation.
Model 1: crude RR. Model 2: adjusted for maternal factors including intended place of birth, parity, age, ethnic background, and neighborhood. Model 3: adjusted for maternal factors and child factors; model 2 + gestational age and presence of Big 4.

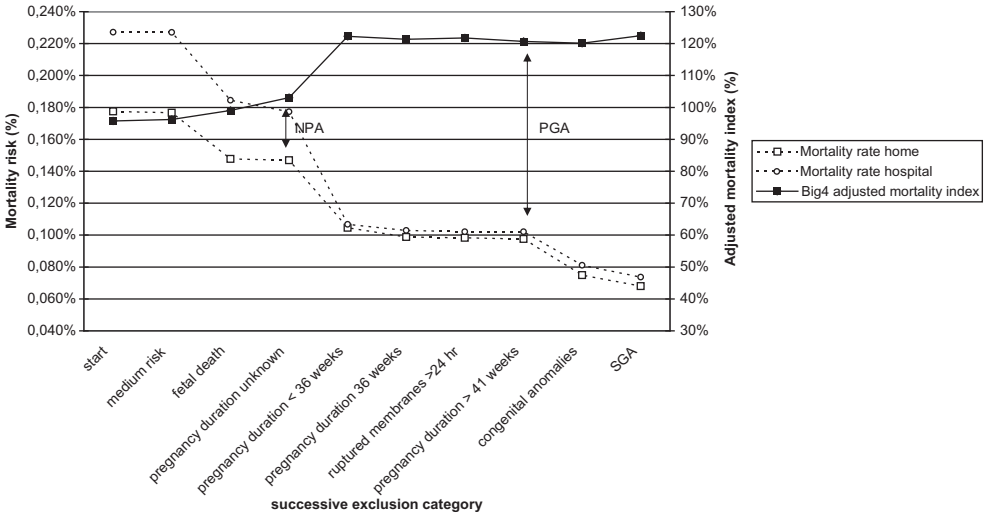


Figure 3.3 Big 4 adjusted mortality index of home birth (hospital based birth under midwife supervision=100%).

pregnancy duration <36 weeks. Combined with the similar crude mortality rates of home and hospital delivery from then onwards, this suggests setting safety for the risk groups still included i.e. all groups right to the exclusion label 'pregnancy duration <36 weeks'. For example after exclusion of pregnancy duration > 41 weeks (PGA group), the adjusted mortality index is 120%, which is slightly larger than the non significant regression result of 111% (table 3.3).

DISCUSSION

Planned home birth within the Dutch maternity care system has a lower crude mortality rate compared to a community midwife led planned hospital birth. However, after case mix adjustment, the relation is reversed, showing a non-significant increased perinatal mortality rate of home birth. Excess setting dependent mortality may arise at home if risk conditions are present or emerge at birth, yet remnant confounding by indication effect (Big4 conditions are more prevalent in hospital) and low mortality prevalence limits statistical proof. Authors favouring a comparison of settings among 'suitable' home births only (PGA), usually exclude risk conditions with a potential setting effect. This mechanism may explain the apparently contradictory results from previous studies.[1,5,7,10-15,17,18]

A strength of this study was the size of the study population, which reflects the complete Dutch experience from 2000-2007. The amount of missing explanatory data is negligible, mortality data have been shown to be complete. No annual trends are observed in the relations shown, except for a minimal gradual decrease in total perinatal mortality.[3]

Our case mix adjustment proved to be essential. The assumption of comparability across home vs. hospital populations appeared not to be justifiable judging from the unequal prevalence of Big4 conditions. These primarily have their origin in early negative fetal conditions and disadvantaged genetic background of the parents. Only in the case of low Apgar, one may argue that the midwifery management during labour might influence it's occurrence, while a management role in SGA, spontaneous prematurity, and congenital anomalies at that stage is unlikely. We decided to include low Apgar cases assuming the role of management to be small compared to the disadvantage of the home setting once a child with persistent low Apgar is born. Thus, our point of departure starts from the risk challenge represented by Big4 at the onset of labour, and investigates whether setting matters in terms of prognosis. The mechanisms underlying the apparent favourable selection for home birth are still to be elucidated. Self selection by the pregnant women can coincide with implicit or explicit selection by the midwife who may tend to 'refer' to hospital if she feels uncomfortable with the risk level at home. The difference in the ratio home:hospital community midwifery led deliveries among the four largest Dutch cities suggests the presence of substantial professional and setting effects. In Amsterdam and Utrecht the ratio is 2:1, and in Rotterdam and the Hague it is 1:2.

Several study limitations merit discussion. While an improvement compared to previous studies, our case mix control is still incomplete because Big4 is unrelated to 15% of deaths. In the PGA population this proportion is even 48%. Thus we cannot rule out remnant confounding by indication as little is known on the factors underlying choice of setting.

RCT would be the superior design to address our research question. However when home birth was part of a trial, participation hampered[24] and introduced selective participation which limited generalisability. Moreover if following one's choice impacts outcome, estimates of setting effects are also biased.[24-26] Despite their shortcomings, in particular when considering the difficulty to overcome the confounding by indication phenomenon, observational studies as ours are therefore invaluable. A comparison with a 100% gynaecologist hospital-based system is not included. The data from an otherwise very similar country as Flanders[27] suggest that more favourable results may be expected in low risk women in general from a hospital-based system. In Flanders perinatal mortality is about 33% less than in the Netherlands, while the caesarean section rates show little difference.

This study primarily focuses upon the disadvantages and neglects the claimed benefits when comparing planned home versus planned hospital births. However studies accessing mother's opinion show that preventing these disadvantages easily outweighs the claimed benefits.[28]

Our results appear compatible with most other reports even though previous studies show conflicting results. Planned home births attended by registered professional attendants are not associated with an increased risk of adverse perinatal outcomes in cohort studies in North America[7,12], the United Kingdom[14], Europe [5,11,15,17], Australia[29] and New Zealand[30]. In contrast, other cohort studies have shown a higher risk of perinatal mortality in planned home births compared to planned hospital births.[10,13,16,18,30] All studies are limited by voluntary submission of data[7,8,11-14,17,31,32], non representative sampling[5,13], lack of appropriate comparison groups[7,12,15,29], or insufficient statistical power[5,17,29,32]. A critical factor, as our study shows, is the in retrospect exclusion of unplanned and unsuitable home births from analysis.[18]

Our results partly agree with those of Kennare at al.[30] who found higher standardised perinatal mortality ratios among planned home deliveries after limited adjustment (birth weight, gestational age). Our results also partly agree with the meta-analysis by Wax et al.[9]: differences in the prevalence of SGA, premature births and congenital anomalies seem equally present in planned home vs. hospital births. They reported a twofold higher neonatal mortality rate but no increase in perinatal mortality. These results are in agreement with figure 3.3 where the fetal death subgroup does not benefit from setting safety. It should be noted that the study of Wax et al. received methodological criticisms[33-36] most notably the inclusion of the study of Pang and the exclusion of the study of De Jonge. Our conclusions apparently contradict those of De Jonge et al. who concluded equal intrapartum and early neonatal outcome of planned home birth vs. hospital birth in apparently the same population.[15] However, the point of departure is not the same. Of our two comparisons of home delivery vs. hospital delivery, one parallels the approach of De Jonge. Our principal approach (NPA) compares neonatal mortality in the actual populations delivering at home vs. hospital, while the approach of De Jonge compares neonatal mortality in a hypothetical group resembling our PGA population. Our adjustment procedure however goes further than the maternal factor adjustment of De Jonge.[15]

From our study we conclude that planned home birth, under routine conditions, is not associated with a higher intrapartum and early neonatal mortality rate. However in subgroups additional risk cannot be excluded.

REFERENCES

1. Home birth--proceed with caution. Lancet. 2010; 376:303.
2. Royal College of Obstetricians and Gynaecologists/ Royal College of Midwives. Home births. Joint statement No.2; 2007 (Online: http://www.rcog.org.uk/files/rcog-corp/uploaded-files/JointStatmentHomeBirths2007.pdf).
3. Ravelli AC, Tromp M, van Huis M, et al. Decreasing perinatal mortality in The Netherlands, 2000-2006: a record linkage study. J Epidemiol Community Health. 2009;63:761-5.
4. CVZ. List of Obstetric Indications (in Dutch: Verloskundig Vademecum) 2003. Diemen (online: http://www.knov.nl/docs/uploads/Verloskundig_Vademecum_2003.pdf).
5. Ackermann-Liebrich U, Voegeli T, Gunter-Witt K, et al. Home versus hospital deliveries: follow up study of matched pairs for procedures and outcome. Zurich Study Team. BMJ. 1996;313:1313-8.
6. Hutton EK, Reitsma AH, Kaufman K. Outcomes associated with planned home and planned hospital births in low-risk women attended by midwives in Ontario, Canada, 2003-2006: a retrospective cohort study. Birth. 2009;36:180-9.
7. Janssen PA, Saxell L, Page LA, Klein MC, Liston RM, Lee SK. Outcomes of planned home birth with registered midwife versus planned hospital birth with midwife or physician. CMAJ. 2009;181: 377-83.
8. Johnson KC, Daviss BA. Outcomes of planned home births with certified professional midwives: large prospective study in North America. BMJ. 2005;330:1416.
9. Wax JR, Lucas FL, Lamont M, Pinette MG, Cartin A, Blackstone J. Maternal and newborn outcomes in planned home birth vs. planned hospital births: a metaanalysis. Am J Obstet Gynaecol. 2010;203:243 e1-8.
10. Wax JR, Pinette MG, Cartin A, Blackstone J. Maternal and newborn morbidity by birth facility among selected United States 2006 low-risk births. Am J Obstet Gynaecol. 2010;202:152 e1-5.
11. Wiegers TA, Keirse MJ, van der Zee J, Berghs GA. Outcome of planned home and planned hospital births in low risk pregnancies: prospective study in midwifery practices in The Netherlands. BMJ. 1996;313:1309-13.
12. Anderson RE, Murphy PA. Outcomes of 11,788 planned home births attended by certified nurse-midwives. A retrospective descriptive study. J Nurse Midwifery. 1995;40:483-92.
13. Bastian H, Keirse MJ, Lancaster PA. Perinatal death associated with planned home birth in Australia: population based study. BMJ. 1998;317:384-8.
14. Chamberlain G, Wraight A, Crowley P. Birth at home. Pract Midwife. 1999;2:35-9.
15. De Jonge A, van der Goes BY, Ravelli AC, et al. Perinatal mortality and morbidity in a nationwide cohort of 529,688 low-risk planned home and hospital births. BJOG. 2009;116:1177-84.
16. Evers AC, Brouwers HA, Hukkelhoven CW, et al. Perinatal mortality and severe morbidity in low and high risk term pregnancies in the Netherlands: prospective cohort study. BMJ. 2010;341:c5639.
17. Lindgren HE, Radestad IJ, Christensson K, Hildingsson IM. Outcome of planned home births compared to hospital births in Sweden between 1992 and 2004. A population-based register study. Acta Obstet Gynaecol Scand. 2008;87:751-9.
18. Pang JW, Heffelfinger JD, Huang GJ, Benedetti TJ, Weiss NS. Outcomes of planned home births in Washington State: 1989-1996. Obstet Gynaecol. 2002;100:253-9.
19. Hildingsson IM, Lindgren HE, Haglund B, Radestad IJ. Characteristics of women giving birth at home in Sweden: a national register study. Am J Obstet Gynaecol. 2006;195:1366-72.
20. Hogberg U. Homebirths in a modern setting - a cautionary tale. Acta Obstet Gynaecol Scand. 2008;87:797-9.
21. Neuhaus W, Piroth C, Kiencke P, Gohring UJ, Mallman P. A psychosocial analysis of women planning birth outside hospital. J Obstet Gynaecol. 2002;22:143-9.
22. Perinatal care in The Netherlands 2006 (in Dutch: Perinatale zorg in Nederland 2006). Utrecht: The Netherlands Perinatal Registry; 2008.
23. Bonsel GJ, Birnie E, Denktas S, Poeran J, Steegers EAP. Dutch report: Lijnen in de Perinatale Sterfte, Signalementstudie Zwangerschap en Geboorte 2010. Rotterdam: Erasmus MC; 2010. (online: http://www.nvk.nl/Nieuws/Dossiers/DossierPerinatalezorg.aspx).

24. Hendrix M, Van Horck M, Moreta D, et al. Why women do not accept randomisation for place of birth: feasibility of a RCT in The Netherlands. BJOG. 2009;116:537-42; discussion 42-4.

25. Tyson H. Outcomes of 1001 midwife-attended home births in Toronto, 1983-1988. Birth. 1991; 18: 14-9.

26. Dowswell T, Thornton JG, Hewison J, et al. Should there be a trial of home versus hospital delivery in the United Kingdom? BMJ. 1996;312:753-7.

27. Cammu H, Martens G, Landuyt J, De Coen K, Defoort P. Perinatale activiteiten in Vlaanderen 2009. Brussel: Studiecentrum voor Perinatale Epidemiologie; 2010.

28. Bijlenga D, Birnie E, Bonsel GJ. Feasibility, Reliability, and Validity of Three Health-State Valuation Methods Using Multiple-Outcome Vignettes on Moderate-Risk Pregnancy at Term. Value Health. 2009;12:821-7.

29. Woodcock HC, Read AW, Bower C, Stanley FJ, Moore DJ. A matched cohort study of planned home and hospital births in Western Australia 1981-1987. Midwifery. 1994;10:125-35.

30. Kennare RM, Keirse MJ, Tucker GR, Chan AC. Planned home and hospital births in South Australia, 1991-2006: differences in outcomes. Med J Aust. 2010;192:76-80.

31. Gulbransen G, Hilton J, McKay L, Cox A. Home birth in New Zealand 1973-93: incidence and mortality. N Z Med J. 1997;110:87-9.

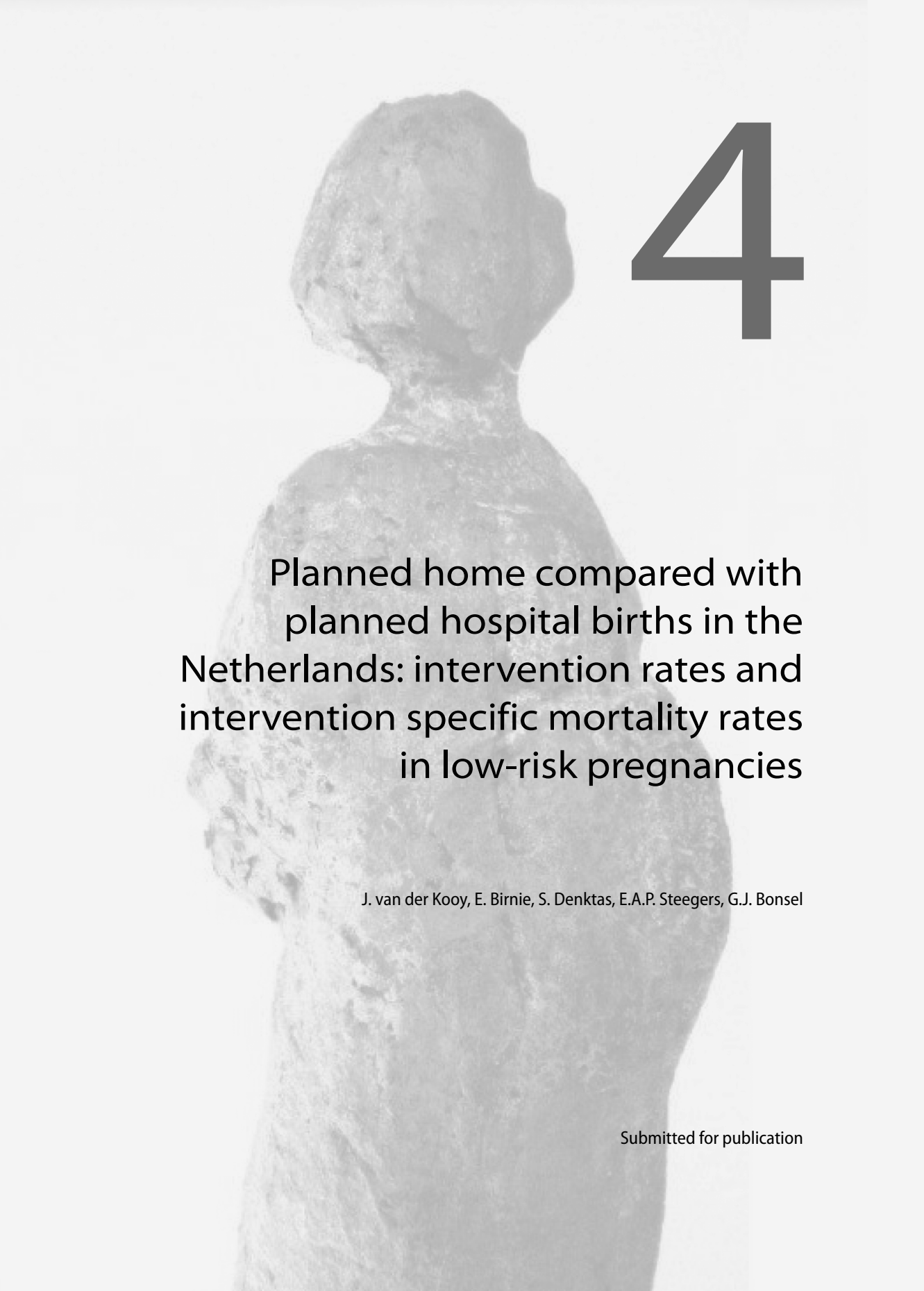
32. Davies J, Hey E, Reid W, Young G. Prospective regional study of planned home births. Home Birth Study Steering Group. BMJ. 1996;313:1302-6.

33. Gyte GM, Dodwell MJ, Macfarlane AJ. Home birth metaanalysis: does it meet AJOG's reporting requirements? Am J Obstet Gynaecol. 2011;204:e15; author reply e8-20, discussion.

34. Johnson KC, Daviss BA. International data demonstrate home birth safety. Am J Obstet Gynaecol. 2011;204:e16-7; author reply e8-20, discussion.

35. Kirby RS, Frost J. Maternal and newborn outcomes in planned home birth vs. planned hospital births: a metaanalysis. Am J Obstet Gynaecol. 2011;204:e16; author reply e8-20, discussion.

36. Sandall J, Bewley S, Newburn M. "Home birth triples the neonatal death rate": public communication of bad science? Am J Obstet Gynaecol. 2011;204:e17-8; author reply e8-20, discussion.



4

Planned home compared with planned hospital births in the Netherlands: intervention rates and intervention specific mortality rates in low-risk pregnancies

J. van der Kooy, E. Birnie, S. Denktas, E.A.P. Steegers, G.J. Bonsel

Submitted for publication

ABSTRACT

Objectives To compare the intervention rate and the intervention specific mortality rates of planned home birth versus hospital birth in community midwife-led deliveries.

Methods The perinatal intervention and mortality rates of 679,952 low-risk women, all offered choice between home and hospital birth, were obtained from the Dutch Perinatal Registry (2000-2007). An intention-to-treat like analysis was performed. Unadjusted and adjusted odds ratios were calculated. Adjustment included maternal, child and health care factors. Child factors (case mix) were based on the presence of at least one of the following risk factors: congenital abnormalities, small for gestational age, preterm birth, or low Apgar score. Moreover, the intervention specific mortality rates by intended place of birth were calculated. The techniques used were nested multiple stepwise logistic regression and stratified analysis.

Results The intervention rate was lower in planned home birth compared to planned hospital births (10.9% 95%CI 10.8-11.0 vs. 13.8% 95% CI 13.6-13.9). Intended place of birth had significant impact on the likelihood to intervene when adjusted for maternal, neonatal (case mix) and health care related factors (planned home birth (OR 0.77 95% CI. 0.75-0.78) versus hospital birth).

The mortality rate was lower in planned home births versus planned hospital births (0.15% vs. 0.18%). When adjusting for maternal, neonatal case mix and health care related factors, the intended place of birth as well as the interaction term with intervention had no significant impact on mortality ($p > 0.05$).

Stratified analysis showed rather similar intervention rates for nulliparous women and universally lower intervention rates for multiparous women planning their delivery at home. For obvious risk groups, e.g. small for gestational age or preterm birth, some disadvantage may exist of lower intervention rates within the planned hospital group.

Conclusion The impact of home birth setting on intervention and perinatal mortality rates differs by sub-risk group. Multiparous women show universally lower intervention rates in planned home births (60%), where in the lowest risk group the presence of under- or overtreatment is difficult to interpret. In primiparous women, the intervention rate is rather similar between home and hospital births. A mortality disadvantage of the setting was observed both within primiparous and multiparous women in the emerging increased risk in a presumed low risk population.

INTRODUCTION

The challenge of obstetric care is to optimize maternal and neonatal health outcomes and the mother's experience of childbirth with the least possible interventions in the normal process.[1] This challenge has led to a widely debate in recent years about relative benefits and risks of birth in different settings and the associated risk of medical interventions.[2-8]

In the Netherlands, approximately 50% of pregnant women start birth in primary care under the supervision of a community midwife. Community midwives are independent health care professionals working either solely or in group practices[9] who provide care for low risk and medium risk pregnant women according to Dutch guidelines.[10] Pregnant women with so called 'medium risk' must give childbirth in the hospital, yet supervised by the community midwife only. Low risk women, on the other hand, can stay with the community midwife and choose the place where to deliver: at home or in the hospital, both supervised by the community midwife only.

The debate on different birth settings in the Netherland has intensified since the national perinatal mortality rate showed to be one of the highest in Europe.[11] While the proportion of home birth deliveries in the Netherlands has steadily decreased to 21% of all births[12], several Anglo-Saxon countries consider the reintroduction of home births. This is based on recent claims of equal safety at lower intervention rates compared to hospital births where overtreatment might be present[13], the claimed reduction of maternal-fetal morbidity and claimed psychosocial advantages for the mother.[2-4,6-8] These benefits may be counterbalanced by the disadvantages associated with delayed treatment or even undertreatment in planned home births leading to an increased risk of perinatal mortality, morbidity and long term adverse effects.[14-15]

Conclusions from previous studies on these claimed benefits and disadvantages can be challenged due to the observational study design; lack of valid case mix adjustment especially when comparing home versus hospital delivery[2-8]; selective exclusion ex post of women who according to the delivery guidelines should have been referred to the gynaecologist before the onset of labour.[2-6]; and failure to calculate intervention specific mortality rates (case fatality analysis).[2-6] These conclusions may therefore have been subject to incomparable groups with the healthiest women opting for a home birth, which may lead to a different tendency to intervene.

The purpose of present study is to compare the intervention rate and the intervention specific mortality rates of home birth versus hospital birth in community midwife-led low-risk deliveries, applying case mix adjustment in an intention-to-treat like approach.

METHODS

Data

The Netherlands Perinatal Registry (PRN) contains population-based information of 96% of all pregnancies in The Netherlands. Source data are collected by 95% of midwives, 99% of gynecologists and 68% of paediatricians (including 100% of Neonatal Intensive Care Unit paediatricians).[11,16] (See www.perinatreg.nl for details.)

Included were the records of all singleton pregnant women (693,592 women) who at the onset of labour were supervised by community midwives between 2000-2007. The onset of labour in the PRN is defined as spontaneous contractions or spontaneous rupture of membranes. Excluded were 13,384 women with so called 'medium risk', e.g. women with a history of postpartum hemorrhage or obesity (BMI > 30) since Dutch guidelines prescribe a hospital delivery for these women. Secondly, we excluded 256 incomplete data records.

The remaining 679,952 women were categorized according to intended place of birth (home/hospital) which usually is concordant with the actual place of birth. For some women the place was undecided or not recorded until the actual onset of labour. In these cases, the intended place was coded 'unknown'.

Outcome measures

Two primary outcomes were defined to assess whether over- or undertreatment is suggested. First, receiving at least one intervention (including operative vaginal delivery and/or caesarean section). Second, perinatal outcome was defined as the intrapartum and early neonatal mortality rate up to 7 days post partum. The PRN does not include long term child outcomes in terms of e.g. psychomotor development and behavioural function.

Case mix adjustment

Detailed information on risk factors is only partially available in the PRN registry. Case mix adjustment is different for mortality and intervention outcomes.

When comparing mortality rates, case mix of any defined group of women was primarily represented by the prevalence (single or combined) of Big4 conditions (see below). The presence of any of the four conditions is known to precede perinatal mortality in 85% of cases (PRN dataset, years 2000-2007, 1.25 million records)[17]. These four conditions were; congenital abnormalities (list defined), intrauterine growth restriction (SGA, birth weight below the 10th percentile for gestational age, gender and parity specific), preterm birth (< 37th week of gestation) or low Apgar score (< 7, measured 5 minutes after birth).[18]

When comparing mortality rates in this paper, we refer to these 4 conditions as the Big4. In the current analysis Big4 represent an objective estimate of the risk load at birth and therefore used as casemix adjustment.

When evaluating intervention rates between planned home and hospital births, the intervention precedes the outcome measure low Apgar score, and should therefore be excluded from the Big4 casemix adjustment. We will continue to refer to this type of casemix adjustment as the Big3. However, when evaluating mortality rates, low Apgar score does precede the outcome mortality, and therefore we included this outcome in the Big4 casemix adjustment. By doing so, we ignore potential differences in policy between planned home and hospital births on the evolvement of these Big4 either antenatally or during delivery. We decided to adjust for low Apgar when comparing mortality, as we assumed the disadvantage of the home setting once a child with low Apgar is born to be more relevant at that stage than the differences in incidence due to differences in setting policy.

Data analysis

As primary analysis we present the results of the intention-to-treat analysis, a nested multiple stepwise logistic regression. The intention-to-treat analysis approach establishes the intervention rates and the intrapartum and early neonatal death rates of planned home versus planned hospital births. It stems from the viewpoint of a pregnant woman starting birth under supervision of a midwife (denominator n=679,952). The intention-to-treat approach thus includes spontaneous preterm labour since to some extent this group was not referred to the gynaecologist at the onset of labour or during labour or was referred late during (home) delivery.

First we compared characteristics of the population by intended place of birth using Student's t-tests for continuous variables with normal distributions and chi-square tests for nominal or ordinal variables. Next we investigated the potential risk of receiving an

intervention (odds ratios) associated with intended place of birth by various predefined nested multivariable logistic regression models (stepwise analysis (likelihood); inclusion $p<0.05$; exclusion $p>0.10$) to which we added maternal, neonatal (case mix) and health care related explanatory variables.

Maternal risk factors were parity (nulliparous/multiparous), age, ethnicity (Western/non-Western; based on a more refined classification in the registry), and living in a deprived neighbourhood (yes/no, based on 4-digit zip-codes and a public list of deprived, zip-code based, neighbourhoods issued by the Dutch government).[19] Health care related factors were time of birth (day 8.00-18.00, night 18.00-8.00), day of birth (week day, weekend) and receiving an intervention (yes/no).

Next we investigated the potential risk of intended place of birth on intrapartum and early neonatal mortality by a set of predefined nested multivariable logistic regression models (stepwise analysis; inclusion $p<0.05$; exclusion $p>0.10$) to which where we added maternal, neonatal (case mix) and health care related explanatory variables respectively (models 2-4). Receiving an intervention and its interaction with intended place of birth was added to the regression model to assess its effect on mortality. For both analyses on mortality and intervention, hospital birth was set reference. All stepwise analyses were repeated with a forward and backward approach, and finally forced inclusion of all predictive variables ($p<0.05$). Risk factor coefficients were only shown if $p<0.05$.

Results across the three approaches were similar unless stated otherwise.

Interpretation of over- and undertreatment

Finally, we calculated intervention specific mortality rates (case fatality analysis) to gain deeper insight when judging the differences in intervention rates in relation with mortality rates, hereby assessing whether undertreatment or overtreatment might be present. In absence of a trial design, we carefully stratified women into predefined risk groups as prior evidence showed that a lower risk (different case mix) is found among planned home births (namely: noBig3, SGA, premature birth, congenital anomaly, combination Big3). Stratified analysis allows for the comparison of mortality rates by receiving an intervention (yes/no) and by planned place of birth (home/hospital). Rate ratios were calculated for the intervention and mortality rates, where the home rate was divided by the hospital rate.

Figure 4.1 displays eight typical intervention patterns that describe the relationship between intervention rate ratio (i.e. intervention rate home/intervention rate hospital), mortality rate

Pattern	Intervention ratio (home/ hospital)	Mortality ratio non intervened (home/hospital)	Mortality ratio intervened (home/hospital)	Selection effects	Suggestive for undertreatment in home	Suggestive for overtreatment in hospital
1	<1 / =	<1 / =	<1 / =	–	–	–
2	<<	<1 / =	<1 / =	+	–	+
3	<1 / =	<1 / =	>1	–	+	–/+
4	<<	<1 / =	>1	+	–/+	+
5	<1 / =	>1	<1 / =	–	+	–
6	<<	>1	<1 / =	+	+	–
7	<1 / =	>1	>1	–	++	–
8	<<	>1	>1	+	++	–

– Very suggestive not to be present.
–/+ Not very suggestive to be present.
+ Suggestive to be present.
++ Very suggestive to be present.

Figure 4.1 Differences in intervention rate and mortality rate between planned home and hospital births classified into eight patterns.

ratios in intervened and non-intervened pregnancies, selection effects and the suggestion of undertreatment in home deliveries or overtreatment in hospital deliveries.

Given this grouping, interpretations whether patterns are suggestive for under- or overtreatment, were made using the mortality ratio in the non intervened and intervened group.

The rate ratios were interpreted as follows:

i. a rate ratio of close or slightly less than 1.0: an equal of slightly lower tendency to intervene in home deliveries in the stratum at hand (</=);
ii. a rate ratio less than 1.0 (typically in the range 0.3-0.6): considerable lower tendency to intervene in home deliveries (<<).

Two effects should be considered when interpreting these patterns. First, residual confounding may still be present, resulting in a favourable case mix in home deliveries. Secondly, selection effects may occur, i.e. a differential tendency to intervene given the planned place of birth. This may select 'healthy' patients receiving an intervention (overtreatment) or 'sick' patients not receiving an intervention (undertreatment).

Suggestive for the presence of undertreatment is when the mortality ratio in the non intervention group exceeds 1, when comparing home vs. hospital. Suggestive for the presence of overtreatment is when selection effects are present and the mortality ratio is about equal in the intervened and non intervened group.

RESULTS

Table 4.1 describes the baseline characteristics of the intention-to-treat-like population (n=679,952).

In the population about 59% of women planned a home delivery and about 32% planned a hospital delivery. Compared to women who planned birth in the hospital or unknown place of birth, women with planned home birth were more likely to be multiparous, 25 years or older, of Dutch origin and to live in a privileged neighbourhood; all conditions for associated with lower intervention rates and lower mortality rates. Neonatal case mix in home birth women also compared favourably. Premature delivery was less common, as was the prevalence of Big4 conditions (home birth 11.0% versus hospital 13.1% versus unknown 12.7%; p<0.001).

Interventions were less common in planned hospital birth (home birth 10.9% versus hospital 13.7% versus unknown 12.2%; p<0.001). Intrapartum and neonatal mortality was 0.15% for planned home births and 0.18% for planned hospital births.

Intervention rates

The crude intervention risk was significantly lower for women who planned home birth (RR 0.76, 95%CI 0.75-0.78, p<0.001) and for women with unknown planned place (RR 0.87, 95%CI 0.84-0.89, p<0.001) compared to those who planned hospital birth (Table 4.2, Model 1). All maternal and neonatal risk factors (except the presence of SGA), showed significant differences in RR in agreement with the expected direction.

The adjusted intervention risks are displayed in models 2-4. Consecutive adjustment for maternal, Big3 case mix and health care related factors showed that the intended place of birth had a significant impact on the likelihood of intervention (planned home birth (OR 0.77, 95%CI. 0.75-0.78; see model 1-3) versus hospital birth. The similarity of crude and adjusted ORs indicates that the differences in tendency to receive an intervention between birth places are unaffected by maternal, Big3 case mix and health care related factors.

Mortality rates

The crude mortality risk was significantly lower for women who planned home birth (RR 0.80 95%CI 0.71-0.91, p<0.001) and for women with unknown intention (RR 0.96 95%CI 0.77-1.19, p<0.001) compared to those who planned hospital birth (Table 4.3, model 1).

Table 4.1 Characteristics and outcome of women in primary care at the onset of labour; intention-to-treat-like approach

	Planned home birth		Planned hospital birth		Planned place unknown	
	n	%	n	%	n	%
Variable	402,912	59%	219,105	32%	57,935	9%
Parity**						
Primiparous	171,986	42.7%	104,249	47.6%	26,254	45.3%
Multiparous (REF)	230,926	57.3%	114,856	52.4%	31,681	54.7%
Maternal Age**						
<19 years	4,036	1.0%	6,713	3.1%	1,190	2.1%
20-25 years	34,661	8.6%	32,617	14.9%	6,823	11.8%
25-34 years (REF)	296,128	73.5%	142,597	65.1%	39,526	68.2%
>35 years	68,087	16.9%	37,178	17.0%	10,396	17.9%
Ethnic background**						
Dutch (REF)	370,647	92.0%	153,572	70.1%	46,966	81.1%
Non Dutch	32,265	8.0%	65,533	29.9%	10,969	18.9%
Neighbourhood**						
Privileged neighbourhood (REF)	388,089	96.3%	196,659	89.8%	53,823	92.9%
Underprivileged neighbourhood	14,823	3.7%	22,446	10.2%	4,112	7.1%
Gestational Age**						
<34wk	2,396	0.6%	1,658	0.8%	567	1.0%
35-36wk	6,510	1.6%	4,064	1.9%	1,206	2.1%
37wk	15,203	3.8%	9,603	4.4%	2,497	4.3%
38-41wk (REF)	372,787	92.5%	200,872	91.7%	52,899	91.3%
>41 wk	6,016	1.5%	2,908	1.3%	766	1.3%

(*continued*)

Table 4.1 Continued

	Planned home birth		Planned hospital birth		Planned place unknown	
	n	%	n	%	n	%
Variable	402,912	59%	219,105	32%	57,935	9%
Big4**						
SGA	28,029	7.0%	18,288	8.3%	4,364	7.5%
Prematurity	8,056	2.0%	5,194	2.4%	1,583	2.7%
Low apgar	1,642	0.4%	1,171	0.5%	290	0.5%
Congenital abnomality	4,711	1.2%	2,826	1.3%	759	1.3%
Combination Big4	1,895	0.5%	1,326	0.6%	373	0.6%
No Big4	358,579	89.0%	190,300	86.9%	50,566	87.3%
Time of delivery**						
Day 8.00 - 18.00 (REF)	167,345	41.5%	96,033	43.8%	24,674	42.6%
Night 18.00 - 8.00	235,567	58.5%	123,072	56.2%	33,261	57.4%
Day of delivery**						
Weekend	109,761	27.2%	59,976	27.4%	15,553	26.8%
Week day (REF)	293,151	72.8%	159,129	72.6%	42,382	73.2%
Interventions**						
Vacuum extraction/forceps	32,481	8.1%	20,404	9.3%	4,630	8.0%
Section cesarean	11,285	2.8%	9,731	4.4%	2,412	4.2%
No vacuum/forceps or section cesarean	359,146	89.1%	188,970	86.2%	50,893	87.8%
Intrapartum & early neonatal death (7days)**						
No	402,266	99.8%	218,672	99.8%	57,826	99.8%
Yes	594	0.15%	403	0.18%	102	0.18%

* Totals may not add up to 100 because of rounding error.
** $p<0.001$ (home vs. hospital vs. unknown planned place of birth).
REF = reference group.

Table 4.2 Intervention (operative vaginal delivery and caesarean section) in women who are in primary care at the onset of labour (intention-to-treat-like approach)

					Model 1			p	Model 2			p	Model 3			p	Model 4			p
	TOTAL(n)	IV (N)	%	p	Crude RR	95%CI			Adj OR	95%CI			Adj OR	95%CI			Adj OR	95%CI		
Intended place of birth				**				**				**				**				**
Home	402,912	43,766	0.109		0.76	0.75	0.78		0.77	0.76	0.78		0.77	0.75	0.78		0.77	0.75	0.78	
Hospital (REF)	219,105	30,135	0.138		1				1				1				1			
Unknown	57,935	7,042	0.122		0.87	0.84	0.89		0.86	0.84	0.89		0.86	0.84	0.89		0.86	0.84	0.89	
Parity				**				**				**				**				**
Primiparous	302,489	70,334	0.233		10.49	10.27	10.71		11.91	11.65	12.17		11.94	11.68	12.20		11.87	11.61	12.13	
Multiparous (REF)	377,463	10,609	0.028		1				1				1				1			
Maternal Age				**				**				**				**				**
<19 years	11,939	1,295	0.108		0.86	0.81	0.91		0.40	0.37	0.42		0.40	0.37	0.42		0.40	0.37	0.42	
20-25 years	74,101	8,737	0.118		0.94	0.92	0.96		0.58	0.57	0.60		0.58	0.57	0.60		0.58	0.57	0.60	
25-34 years (REF)	478,251	59,536	0.124		1				1				1				1			
>35 years	115,661	11,375	0.098		0.77	0.75	0.78		1.37	1.34	1.40		1.36	1.33	1.39		1.36	1.33	1.40	
Ethnic background				**				**				**				**				**
Dutch (REF)	571,185	69,983	0.123		1				1				1				1			
Non Dutch	108,767	10,960	0.101		0.80	0.79	0.82		0.96	0.94	0.98		0.96	0.94	0.99		0.96	0.94	0.99	
Neighbourhood				**				**				**				**				**
Privileged neighbourhood (REF)	638,571	76,646	0.120		1				1				1				1			
Underprivileged neighbourhood	41,381	4,297	0.104		0.85	0.82	0.88		0.90	0.87	0.94		0.90	0.87	0.93		0.90	0.87	0.93	
Gestational Age				**				**								**				**
<34wk	4,621	873	0.189		1.90	1.76	2.04						1.02	0.94	1.11		1.02	0.94	1.11	
35-36wk	11,780	1,926	0.163		1.48	1.41	1.55						0.94	0.90	1.00		0.94	0.89	0.99	
37wk	27,303	2,701	0.099		0.83	0.80	0.86						0.64	0.61	0.66		0.63	0.61	0.66	
38-41wk (REF)	626,558	73,286	0.117		1								1				1			
>41 wk	9,690	2,157	0.223		2.16	2.06	2.27						2.09	1.98	2.21		2.10	1.99	2.21	

(*continued*)

Table 4.2 Continued

					Model 1			p	Model 2			p	Model 3			p	Model 4			p
	TOTAL(n)	IV (N)	%	p	Crude RR	95%CI			Adj OR	95%CI			Adj OR	95%CI			Adj OR	95%CI		
Big3				**				**								**				**
SGA	50,681	5,169	0.102		0.83	0.81	0.85						0.86	0.83	0.88		0.86	0.83	0.89	
Prematurity	16,401	2,799	0.350																	
Congenital abnomality	3,594	904	0.252		1.86	1.76	1.97						1.80	1.70	1.91		1.80	1.69	1.91	
Combination Big3	80,507	11,238	0.140		2.64	2.44	2.85						2.28	2.09	2.48		2.28	2.09	2.48	
No Big3	533,771	59,178	0.111		1								1				1			
Time of delivery								**												**
Day 8.00 - 18.00 (REF)	288,052	38,414	0.133		1												1			
Night 18.00 - 8.00	391,900	42,529	0.109		0.79	0.78	0.80										0.86	0.84	0.87	
Day of delivery								**												
Week day (REF)	494,662	58,785	0.119		1												nie			
Weekend	185,290	22,158	0.120		1.01	0.99	1.02													

Model 1: crude RR.
Model 2: adjusted for motherfactors.
Model 3: adjusted for motherfactors + child factors.
Model 4: adjusted for motherfactors + child factors + health care factors.
* p<0.05 (categories of variables).
** p<0.001(categories of variables).
REF = Reference.
nie = not in equation.

All maternal and neonatal risk factors, except the presence of a single SGA, showed significant differences in RR in agreement with the expected direction.

When adjusting for maternal, Big4 case mix and health care related factors respectively, the intended home birth group who received an intervention had a significant impact on the likelihood of mortality compared to the intended hospital birth group receiving no intervention, where as the intended hospital group who received intervention did not have a significant impact.

Table 4.4 describes the results of the case fatality analysis for certain risk groups. High mortality rates were seen within in the primiparous with a Big3 combination (6.8%, data not shown) and multiparous women (9.2%, data not shown). The intervention rate was lower for women who planned to give birth at home, except for congenital anomalies in primiparous women (RR=1.03).

Grouping the intervention rate ratios into various risk groups, we observed the following patterns as observed in table 4. 4.

The primiparous NoBig3 (pattern 1) accounts for 39% of all deliveries, and the multiparous NoBig3 group (pattern 4) accounts for 50% of all deliveries, the remaining risk groups for 11%.

DISCUSSION

The intervention rate of planned home birth in the Dutch maternity care system is lower compared to a community midwife led planned hospital birth after case mix adjustment, especially in multiparous women. Although we cannot provide an overall judgement on the impact of birth setting on the presence of under- and overtreatment, this seems justified for some risk groups. In the NoBig3 primiparous group equal intervention rates seem present. The surprisingly low mortality rate in this primiparous group might be an artefact due to lack of complete case mix adjustment or a true advantage of avoiding hospital-related risks. In the Big3 primiparous group, the mortality rate was lower in planned hospital births than in planned home births, suggestive for undertreatment in home births in these emerging risk groups at birth.

In multiparous women, in contrast, the substantially lower intervention rate associated with planned home births and its association with under- and overtreatment is much more diverse. In Big3 multiparous groups some disadvantage of undertreatment exists. In

Table 4.3 Intrapartum and neonatal death 0-7 days in women who are in primary care at the onset of labour (intention-to-treat-like approach)

					Model 1			p	Model 2			p	Model 3			p
	TOTAL(n)	mortality (n)	%	p	Crude RR	95%CI			Adj OR	95%CI			Adj OR	95%CI		
Parity				**				**					nie			
Primiparous	302,489	667	0.22%		1.58	1.40	1.78		1.01	0.87	1.16					
Multiparous (REF)	377,463	521	0.14%		1				1							
Maternal Age				**				**				**				**
<19 years	11,939	44	0.37%		2.43	1.78	3.32		1.77	1.24	2.52		1.79	1.26	2.54	
20-25 years	74,101	146	0.20%		1.24	1.03	1.49		0.98	0.80	1.21		0.99	0.80	1.21	
25-34 years (REF)	478,251	753	0.16%		1				1				1			
>35 years	115,661	245	0.21%		1.38	1.19	1.60		1.50	1.28	1.77		1.50	1.28	1.76	
Ethnic background				**				**				**				**
Dutch (REF)	571,185	953	0.17%		1				1				1			
Non Dutch	108,767	235	0.22%		1.31	1.13	1.52		1.27	1.06	1.51		1.29	1.09	1.53	
Neighbourhood												*				**
Privileged neighbourhood (REF)	638,571	1,113	0.17%		1				1				1.00			
Underprivileged neighbourhood	41,381	75	0.18%		1.02	0.80	1.30		0.70	0.53	0.92		0.70	0.53	0.93	
Gestational Age				**				**				**				**
<34wk	4,621	378	8.18%		87.01	76.10	99.48		27.79	22.62	34.15		27.71	22.58	34.01	
35-36wk	11,780	72	0.61%		5.73	4.43	7.41		2.34	1.74	3.15		2.35	1.75	3.17	
37wk	27,303	61	0.22%		2.12	1.61	2.79		2.05	1.55	2.70		2.04	1.55	2.69	
38-41wk (REF)	626,558	663	0.11%		1				1				1			
>41 wk	9,690	14	0.14%		1.39	0.80	2.40		1.13	0.65	1.97		1.15	0.66	2.00	

Big4				**							**			**
SGA	50,681	81	0.16%	0.96	0.76	1.21		2.45	1.92	3.12	2.45	1.92	3.12	
Low apgar	31.03	111	3.58%	22.02	17.84	27.18	**	45.78	36.45	57.50	47.56	37.95	59.61	
Congenital abnomality	8,296	90	1.08%	5.46	4.28	6.97	**	12.31	9.52	15.92	12.46	9.64	16.10	
Combination Big4	3,594	420	11.69%	115.01	101.14	130.77	**	35.59	29.17	43.42	35.96	29.46	43.90	
No Big4	599,445	309	0.05%	1				1			1			
Time of delivery											nie	nie	nie	
Day 8.00 - 18.00 (REF)	288,052	499	0.17%	1				1						
Night 18.00 - 8.00	391,900	689	0.18%	1.02	0.91	1.16		1.09	0.96	1.24				
Day of delivery											nie	nie	nie	
Week day (REF)	494,662	841	0.17%	1				1						
Weekend	185,290	347	0.19%	1.12	0.98	1.28		1.10	0.95	1.26				
Intervention*intended place of birth														
Home intervention				2.14	1.80	2.56		1.50	1.19	1.88	1.51	1.25	1.84	
Home no intervention				0.63	0.56	0.71		0.95	0.81	1.11	nie			
Hospital intervention				2.07	1.68	2.55		1.38	1.08	1.78	nie			
Hospital no intervention (REF)				1				1			1			
Unknown intervention				1.77	1.14	2.76		0.64	0.38	1.09	nie			
Unknown no intervention				1.00	0.80	1.25		0.63	0.45	0.86	nie			

Model 1: crude RR.
Model 2: forced analysis.
Model 3: stepwise analysis.
* p<0.05 (categories of variables).
** p<0.001(categories of variables).
REF = Reference.
nie=not in equation.

Table 4.4 Intervention rate (operative vaginal delivery and caesarean section) and mortality (intrapartum and neonatal death 0-7days post partum) subdivided into riskgroups and place of delivery

				No Intervention Group						Intervention Group								
			Total N	N		Mortality*		Mortality Rate		N		Mortality*		Mortality Rate		Intervention Ratio**	Mortality Ratio No Intervention Group***	Mortality Ratio Intervention Group****
Pattern			Home + Hospital	Home	Hospital	Home	Hospital	Home	Hospital	Home	Hospital	Home	Hospital	Home	Hospital	Home vs. Hospital	Home vs. Hospital	Home vs. Hospital
1	Nobig3	P0	242,184	116,663	68,641	123	78	0.11%	0.11%	35,179	21,701	63	52	0.18%	0.24%	0.95	0.93	0.75
4		P1	310,398	205,324	96,889	127	74	0.06%	0.08%	3,582	4,603	27	14	0.75%	0.30%	0.37	0.81	2.48
3	Sga	P0	20,071	9,186	6,799	16	13	0.17%	0.19%	2,335	1,751	13	5	0.56%	0.29%	0.99	0.91	1.95
8		P1	26,187	16,188	9,377	16	7	0.10%	0.07%	281	341	3	1	1.07%	0.29%	0.48	1.32	3.64
3	Premature Birth	P0	8,752	4,366	2,582	41	36	0.94%	1.39%	1,088	716	6	1	0.55%	0.14%	0.90	0.67	3.95
2		P1	4,278	2,339	1,592	18	30	0.77%	1.88%	152	195	3	4	1.97%	2.05%	0.53	0.41	0.96
7	Congenital Anomaly	P0	3,785	1,511	957	12	6	0.79%	0.63%	815	502	9	5	1.10%	1.00%	1.03	1.27	1.11
4		P1	3,740	2,300	1,253	20	11	0.87%	0.88%	78	109	1	1	1.28%	0.92%	0.39	0.99	1.40
1	Combination Big3	P0	1,134	487	318	34	23	6.98%	7.23%	186	143	10	10	5.38%	6.99%	0.85	0.97	0.77
8		P1	912	487	314	46	29	9.45%	9.24%	54	57	6	3	11.11%	5.26%	0.61	1.02	2.11
Column			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16

* mortality is defined as intrapartum and neonatal death (0-7 days post partum).

** number iv group (home) / number non iv group (home) divided by number iv group (hospital) / number non iv group (hospital).

*** mortality iv group (home) / total number of people iv group (home) divided by mortality iv group (hospital) / total number of people iv group (hospital).

**** mortality non iv group (home) / total number of people non iv group (home) divided by mortality non iv group (hospital) / total number of people non iv group (hospital).

the NoBig3 multiparous group (about 50% of total), however, the benefit of substantially fewer interventions in the NoBig3 multiparous home group seems to be counterbalanced by substantially increased mortality if intervention occurs. Observational data cannot demonstrate that the outcome of these interventions would have profited from a planned hospital birth, since data point to good risk selection here as the mortality rate in the no intervention multiparous noBig3 group of planned home births is extremely low. If risk selection can be improved both in terms of detection and timely referral, in particular multiparous women could experience benefits from the non-medical setting at no price.

We found the mortality rate in Big3 pregnancies generally less favourable in planned home births compared to planned hospital births. Possible explanations are overtreatment in the planned hospital group, selection of only worst cases within the home group ('undertreatment'), or delayed timing of referral. One important source of delay is travel time from home to the hospital. Amelink et al found that 0.4% of all low risk pregnancies need urgent referral. In the Netherlands, average time to the nearest hospital is about 13 minutes (ranging from 0 to 60 minutes), but about 30 minutes should be added for ambulance arrival and patient preparation. They concluded that the net travel time from home to hospital of 20 minutes or more by car is associated with an increased risk of mortality and adverse outcomes in term women.[20] Moreover, Ravelli et al found that delivery at 37 weeks of gestation or 41 weeks of gestation in combination with travelling time increased the risk of mortality even further.[21] A second source of delay is the delay of the referral decision as suggested by Evers.[15] They observed a more than 3.5-fold higher perinatal death rate in women who were referred from primary to secondary care during labour compared with infants of women who started labour in secondary care.

Strengths and limitations

A strength of this study was the size and completeness of the study population, covering the complete Dutch experience from 2000-2007. The amount of missing data was negligible and mortality data have been shown to be complete. Annual trends in the studied relations were absent, except for a minimal gradual decrease in total perinatal mortality[11].

Our case mix adjustment turned out to be essential. We previously showed that, within the low risk group of midwife led deliveries, unequal prevalence of Big4 conditions is present in planned home versus hospital births. This suggests an unequal risk load at the onset of

childbirth since, either due to self selection or due to the midwife's proposal, the healthiest and most affluent women are more likely to receive home birth. When one fails to adjust for this, one may introduce confounding by indication bias.[18]

Another strength was the inclusion of women ex post, who according to delivery guidelines should have been referred to the gynaecologist before the onset of labour. We previously showed, that under the Dutch system, health care performance during labour should include the performance of the preliminary antenatal phase in terms of distinguishing between low risk (midwife) and high risk (gynaecologists) pregnancies (intention-to-treat-analysis like approach).[18]

Several limitations merit discussion. A RCT would be the superior design to address our research question, in particular as the referral process and the indication for an intervention interact, and are subject to external effects (e.g. travel time, judgement of caregiver) which may differ by intended place of birth. However the only RCT on home versus hospital birth resulted in low participation rates and introduced selective participation.[22] Treatment groups composed on the basis of women's preference for setting is likely to affect outcome, producing biased estimates of setting effects. Hence a RCT design is unfeasible within our obstetric system[22-24]. As next-best option, we applied case mix adjustment to the extent the data permitted. A severe limitation is that few data are available on the precise clinical assessment leading to referral or intervention which would allow for better judgement on setting-dependent overtreatment or undertreatment in our analysis. Moreover, the Big4 adjustment does not adjust for potential differences in morbidity associated with the remaining 15% of perinatal deaths not covered by Big4.

We are aware that a comparison with low risk women planning a gynecologist-led hospital birth is not included as this option is unavailable in the Dutch system. An observational study by Maassen et al on gynaecologist led care of presumably low risk women, is difficult to interpret, since detailed risk factors are not routinely collected, thereby limiting case mix adjustment.[25]

Our study is limited in that only intervention rate and mortality are used as outcome indicators, ignoring mother's experience. However, studies addressing the trade-off between intervention disadvantages to the mother (e.g. caesarean section) versus safety of the child clearly indicate that even small advantages to the child's outcome outweigh the disadvantage of an intervention or the general disadvantage of birth in a hospital[26].

Our results appear compatible with most of the few available reports on this issue. Previous studies on planned home births attended by registered community midwives confirm the lower risk of receiving an intervention and suggest equal mortality.[2-8] However these

studies are limited by lack of applying complete case mix adjustment, thereby suggesting risk equivalence of home and hospital groups[2-8], ex post exclusion of unplanned and unsuitable home births from analysis.[2-6,8], voluntary submission of data[4-6,8], or lack of statistical power[2-3,5,8]. These limitations generally tend to benefit outcome in favour of home birth. Surprisingly, none of these studies has performed a case fatality analysis based on predefined risk groups.

While multiparous women show universally lower intervention rates (60%) in planned home births, interpretation whether undertreatment in home births or overtreatment in hospital births is present, is difficult in the low risk group (NoBig3). However, in emerging increased risk in a presumed low risk groups (Big3) undertreatment in planned home birth seems present. In primiparous women, the intervention rate is rather similar for planned place of birth. A similar mortality disadvantage of setting suggestive for undertreatment in home deliveries was observed in emerging risk groups (Big3).

REFERENCES

1. World Health Organisation MaNHCSMU. Care in Normal Birth: A practical guide. Geneva. Switzerland. WHO/TRH/MSM/9624. 1996.
2. Ackermann-Liebrich U, Voegeli T, Gunter-Witt K, et al. Home versus hospital deliveries: follow up study of matched pairs for procedures and outcome. Zurich Study Team. Bmj. 1996 Nov 23;313(7068):1313-8.
3. Hutton EK, Reitsma AH, Kaufman K. Outcomes associated with planned home and planned hospital births in low-risk women attended by midwives in Ontario, Canada, 2003-2006: a retrospective cohort study. Birth. 2009 Sep;36(3):180-9.
4. Janssen PA, Saxell L, Page LA, Klein MC, Liston RM, Lee SK. Outcomes of planned home birth with registered midwife versus planned hospital birth with midwife or physician. Cmaj. 2009 Sep 15;181(6-7):377-83.
5. Johnson KC, Daviss BA. Outcomes of planned home births with certified professional midwives: large prospective study in North America. Bmj. 2005 Jun 18;330(7505):1416.
6. Lindgren HE, Radestad IJ, Christensson K, Hildingsson IM. Outcome of planned home births compared to hospital births in Sweden between 1992 and 2004. A population-based register study. Acta Obstet Gynecol Scand. 2008;87(7):751-9.
7. Olsen O. Meta-analysis of the safety of home birth. Birth. 1997 Mar;24(1):4-13; discussion 4-6.
8. Wiegers TA, Keirse MJ, van der Zee J, Berghs GA. Outcome of planned home and planned hospital births in low risk pregnancies: prospective study in midwifery practices in The Netherlands. Bmj. 1996 Nov 23;313(7068):1309-13.
9. Amelink-Verburg MP, Buitendijk SE. Pregnancy and labour in the Dutch maternity care system: what is normal? The role division between midwives and obstetricians. J Midwifery Womens Health. 2010 May-Jun;55(3):216-25.
10. CVZ. CV. List of Obstetric Indications (in Dutch: Verloskundig Vademecum) 2003. Diemen (http://www.knov.nl/docs/uploads/Verloskundig_Vademecum_2003.pdf). 2003.
11. Ravelli AC, Tromp M, van Huis M, et al. Decreasing perinatal mortality in The Netherlands, 2000-2006: a record linkage study. J Epidemiol Community Health. 2009 Sep;63(9):761-5.
12. Nederland SPR. Perinatale Zorg in Nederland 2007. 2009.
13. Royal College of Obstetricians and Gynaecologists/ Royal College of Midwives. Home births. Joint statement No.2; 2007 (online: http://www.rcog.org.uk/files/rcog-corp/uploaded-files/JointStatmentHomeBirths2007.pdf).

14. Mori R, Dougherty M, Whittle M. An estimation of intrapartum-related perinatal mortality rates for booked home births in England and Wales between 1994 and 2003. Bjog. 2008 Apr;115(5):554-9.

15. Evers AC, Brouwers HA, Hukkelhoven CW, et al. Perinatal mortality and severe morbidity in low and high risk term pregnancies in the Netherlands: prospective cohort study. Bmj. 2010;341:c5639.

16. Perinatal care in The Netherlands 2006 (in Dutch: Perinatale zorg in Nederland 2006). Utrecht: The Netherlands Perinatal Registry; 2008.

17. Bonsel GJ, Birnie E, Denktas S, Poeran J, Steegers EAP. Dutch report: Lijnen in de Perinatale Sterfte, Signalementstudie Zwangerschap en Geboorte 2010 (http://www.nvk.nl/Nieuws/Dossiers/DossierPerinatalezorg.aspx). Rotterdam: Erasmus MC; 2010.

18. van der Kooy J, Poeran J, de Graaf JP, et al. Planned home compared with planned hospital births in the Netherlands: intrapartum and early neonatal death in low-risk pregnancies. Obstet Gynecol. 2011 Nov;118(5):1037-46.

19. The MolatEVGdodnatpScftdodn, Hague: VROM IDAoawrnd-e-pkb-k-w-o-r-v-a-e-p-p.

20. Amelink-Verburg MP, Verloove-Vanhorick SP, Hakkenberg RM, Veldhuijzen IM, Bennebroek Gravenhorst J, Buitendijk SE. Evaluation of 280,000 cases in Dutch midwifery practices: a descriptive study. Bjog. 2008 Apr;115(5):570-8.

21. Ravelli AC, Jager KJ, de Groot MH, et al. Travel time from home to hospital and adverse perinatal outcomes in women at term in the Netherlands. Bjog. 2011 Mar;118(4):457-65.

22. Hendrix M, Van Horck M, Moreta D, et al. Why women do not accept randomisation for place of birth: feasibility of a RCT in The Netherlands. Bjog. 2009 Mar;116(4):537-42; discussion 42-4.

23. Tyson H. Outcomes of 1001 midwife-attended home births in Toronto, 1983-1988. Birth. 1991 Mar;18(1):14-9.

24. Dowswell T, Thornton JG, Hewison J, et al. Should there be a trial of home versus hospital delivery in the United Kingdom? Bmj. 1996 Mar 23;312(7033):753-7.

25. Maassen MS, Hendrix MJ, Van Vugt HC, Veersema S, Smits F, Nijhuis JG. Operative deliveries in low-risk pregnancies in The Netherlands: primary versus secondary care. Birth. 2008 Dec;35(4):277-82.

26. Bijlenga D, Birnie E, Bonsel GJ. Feasibility, Reliability, and Validity of Three Health-State Valuation Methods Using Multiple-Outcome Vignettes on Moderate-Risk Pregnancy at Term. Value Health. 2009 Mar 10.

PART II

How can Dutch perinatal care be improved, in particular, are there innovative strategies available to improve risk selection and midwife-led birth care?

5

Different settings of place of midwife-led birth: evaluation of a midwife-led birth centre

J. van der Kooy, J.P. de Graaf, E. Birnie, S. Denktas, E.A.P Steegers, G.J. Bonsel

Submitted for publication

ABSTRACT

Background In perinatal care concerns are raised whether the small additional risk of home deliveries is of fact by the claimed advantage of lower intervention rates compared to planned hospital births. Homelike birth centres have been proposed as an alternative.

Objectives This paper addresses whether the introduction of a midwife-led birth centre adjacent to an academic hospital leads to better outcomes.

Method Anonymized data, between January 2007 and June 2012, was collected from participating practices. Women (n=5,558) were categorized according to intended place of birth. Women's characteristics and outcomes were compared between the period before and after its introduction using chi-square and Fisher's Exact tests. Direct and indirect standardized rates were calculated for different outcomes (I) intrapartum and neonatal mortality (7 days), (II) composite outcome of neonatal morbidities, (III) composite outcome of maternal morbidities, and (IV) medical intervention.

Results Women's characteristics were most unfavourable for intended birth centre births. After its introduction neonatal morbidities decreased (5.0% vs. 3.8%) as did maternal morbidities (8.3% vs. 7.3%). Interventions were about equal. Neonatal morbidities occurred more in birth centre births compared to home or hospital births (5.3% vs. 1.9% vs. 2.7% respectively), while maternal morbidities were about equal.

Conclusion Neonatal morbidity and maternal morbidity tended to decrease, while overall intervention rates were unaffected. This change could be interpreted by the redistribution of the higher risk women among the low risk population intending birth at the birth centre instead of home. Alternative explanations are still to be explored.

INTRODUCTION

There are considerable variations in organization of Perinatal Care. In the Netherlands, approximately 50% of women start delivery under supervision of a community midwife. Dutch community midwives are independent health care professionals who provide care for low risk and medium risk pregnant women. Dutch guidelines define low, medium and high risk pregnant women.[1] A low risk pregnant woman who becomes high risk is referred antenatally or during delivery to the gynaecologist for remaining ante- and intrapartum care. Low risk women are allowed to choose the place where to deliver: at home, in the hospital or in a birth centre, all supervised by the community midwife. The frequency of these different midwife-led birth places differs across regions, with fewer home deliveries in urban areas.[2] Pregnant women with so called medium risks should deliver in the hospital according to Dutch guidelines, yet supervised by the midwife only.

One of the concerns raised regarding the Dutch Perinatal Care System is whether the additional risk on adverse perinatal outcome of home deliveries is of fact by the claimed advantage of lower intervention rates compared to planned hospital births.[3-9] Birth centres adjacent to hospitals have been proposed a new setting that combines the advantages of home and hospital.[10-13] They are designed to provide an intermediate option of care between home and hospital birth for low risk women. Despite existing organization differences, birth centre care generally includes a homelike, nonclinical environment, a rather autonomous midwifery practice, and a commitment to and belief in normal, physiologic birth.[11,12,14]

In this paper we address whether the introduction of a midwife-led birth centre adjacent to the hospital combines the advantages of home and hospital deliveries by studying the regional perinatal outcomes before and after the introduction of the birth centre hereby comparing the different places of birth. Additionally, we investigate whether the introduction of a midwife-led birth centre leads to a different risk selection of women planning their delivery either at home, at the hospital or at the birth centre, resulting in an altered risk for perinatal, maternal morbidities and intervention pattern.

METHODS

Birth centre

The birth centre Sophia started care in October 2009. It is a separate unit, consisting of four single birthing rooms and 12 rooms for post partum maternity care. It is located on the same floor as the obstetric labour ward (100 meters), yet with its own entrance and

home-like interior. The unit is staffed by local community midwives. The unit fits to the UK National Perinatal Epidemiology Unit report description of a 'midwife-led birth centre adjacent to the hospital', i.e. "An alongside midwifery care unit offering care to women with straightforward pregnancies during labour and birth in which midwives take primary professional responsibility for care. During labour and birth diagnostic and treatment medical services, including obstetric, neonatal and anaesthetic care are available, should they be needed, in the same building, or in a separate building on the same site. Transfer will normally be by trolley, bed or wheelchair. "[14]

The aim of the unit is to provide a safe, homelike environment where low risk women can give birth. Women with a medical indication requiring a hospital birth under supervision of the midwife are thus excluded (medium risk; a small (1%) category according to current guidelines). The unit aims to provide risk led care, including assessing the risk status of each women at different time intervals and act promptly and accordingly using standardised protocols and the Dutch guidelines.[1] Special attention is given to certain ethnic minority groups and women with a low social economic background both in terms of risk monitoring and tailored provision of care, both covered in standardized protocols. An expert group consisting of a gynaecologist, midwives, maternity nurses, a public health expert and administrating staff is responsible for the continuous development, revision and extension of these protocols and new interventions. Furthermore, the expert group has a standard obstetrician-led audit procedure for all cases with poor intrapartum related outcome.

Local community midwives take full responsibility for the care delivered, thus developing and maintaining their competence. Labour is managed traditionally; the fetal heart rate is monitored with a hand held Doppler apparatus, and interventions are minimal. During labour continuous one-to-one care is given by a maternity nurse. If (acute) complications arise, obstetrical and neonatal expertise and clinical facilities are directly available. Rarely, in non acute cases, transferral from the birth centre to the adjacent hospital is impossible for logistic reasons. Women are then referred to another hospital in Rotterdam. This occurred in 50 (3%) of the cases.

Design and Data

A regional study in the north of Rotterdam was designed comparing outcomes before and after the introduction of the midwife-led birth centre in October 2009, using direct and indirect standardization. Anonymized data were collected from the registries of the four largest local community midwife practices in the region, whose women are allowed to choose

place where to deliver (in the birth centre, at home or in nearby hospitals). If women deliver in the birth centre, the data registration is still the midwife's responsibility through internet entry. Data were collected between January 2007 and June 2012. We selected the records of all singleton pregnancies supervised by a community midwife at the onset of labour (5,953 women) according to their planned place of birth. This also included those pregnant women who according to current guidelines should not deliver under the supervision of the midwife, e.g. preterm deliveries, small for gestational age. When women were preterm at the onset of labour they were often directly referred to secondary care, except for those women who were occasional rapid and too late for transport. The onset of labour was defined as spontaneous contractions or the spontaneous rupture of membranes. Excluded were 62 women with so called 'medium risks' (e.g. women with a history of postpartum haemorrhage or obesity (BMI>35) and another 333 women since their planned place of birth was unknown. The remaining 5,558 were divided into two groups, the period before the introduction of the birth centre (n=1834) and the period after its introduction (n=3724).

Within each period women were categorized by intended place of birth (at home, at the hospital or at the birth centre), which usually is concordant with the true place of birth.

The retrospective use of anonymized medical records exempted institutional review of the Medical Ethics Committee.

Outcomes

Four primary outcomes were chosen: (I) intrapartum and early neonatal mortality (up to 7 days), (II) a composite outcome of intrapartum related neonatal morbidities (neonatal encephalopathy, brachial plexus injury, fractured clavicle, cephaloheamatoma, neonatal infection, low Apgar score (<7 after 5 minutes), neonatal hospital admission, or other trauma related to birth), (III) a composite outcome of intrapartum related maternal morbidities (third/fourth degree rupture or postpartum haemorrhage > 1000cc), and (IV) the presence of a medical intervention (vacuum extraction, forceps extraction, or caesarean delivery). For the latest outcome we also provided results on caesarean delivery only.

Data handling

Of the 27 selected variables, the variable education level had ≥30% missing values and was therefore excluded from analysis. For all other variables missing values were <30% and were replaced by mean, median and mode, for respectively, numeric (normally distributed), ordinal and nominal values.

Data analysis

First, we compared women's characteristics between the period before and after the introduction of the midwife-led birth centre, using chi-square tests and Fisher's Exact tests. We compared women's profiles before and after the introduction of the birth centre, according to the intended place of birth (see table 5.1, column A; home before vs. after the introduction of the birth centre, and B; hospital before vs. after the introduction of the birth centre). Secondly, we compared women's profiles for the different places of birth, after the introduction of the Sophia birth centre (see table 5.1, column C; home vs. birth centre and D; hospital vs. birth centre).

We calculated the detailed birth weight distribution (according to the national birth weight reference curves[15]) for all low risk singleton pregnant women, under supervision of a community midwife at the onset of labour between 2000-2007 in the municipality of Rotterdam (29,357 women), according to the planned place of birth. This served as reference for the birth weight distribution in the four practices, which we calculated for the periods before (1,834 women) and after (3,724 women) the introduction of the birth centre, again according to their intended place of birth.

Finally, since no detailed risk factors are available in the registry, we provided the prevalence of Big3 pregnancies as a proxy for risk load, which can be used to adjust for case mix differences. The term 'Big' was chosen since these conditions were found to be three big causes preceding perinatal mortality. Big3 pregnancies are defined as: congenital abnormalities (list defined), intrauterine growth restriction (SGA, birth weight below the 10th percentile for gestational age, gender and parity specific), or preterm birth (<37th week of gestation).[9] Detailed analysis of the complete perinatal dataset of the Perinatal Registry of the Netherlands (PRN), covering all pregnancies of the years 2000-2007 (1.25 million records), show that the presence of any of these three conditions preceded perinatal mortality in 80% of cases.[16] A p-value (two sided) < 0.05 was considered a statistically significant difference.

The data were analysed with Statistical Package of Social Sciences version 20.0 for Windows (SPSS Inc, Chicago, IL, USA).

Standardization

To increase the validity of our results of the comparison before and after the introduction of the birth centre, we applied direct as well as indirect standardization. The index population (i.e. the population of interest) consisted of the women who could plan their birth the

midwife-led birth centre (n=3,724 of which 470 home, 1,583 hospital and 1,671 birth centre). The standard or reference population consisted of 1,834 eligible women who could not plan their birth in the birth centre (443 planned home births, and 1,391 planned hospital births).

In our analysis the index population was standardized according to the strata of the standard population. The index populations of planned birth centre births and planned hospital births were standardized using the standard populations of planned hospital births, since women's profiles were similar. The index population of planned home birth was standardized using the standard population of planned home births.

This analysis was repeated using the index population which consisted of women who could plan their birth at the midwife-led birth centre (n=3,724 of which 470 home, 1,583 hospital and 1,671 birth centre). Herewith we controlled for historic trends. Within this analysis the planned hospital birth population was set as reference population (n=1,583).

Populations were stratified for parity (nulliparous vs. multiparous), age ($\leq$24 years, 25-34 years, $\geq$35 years), ethnicity (Dutch vs. non-Dutch) and the presence of Small for Gestation Age (SGA, birth weight below the 10th percentile for gestational age, gender and parity specific; yes vs. no). In our analysis the presence of SGA represented an objective estimate of the risk challenge at birth.[9,16]

The direct standardized rates were estimated as a weighted average of the index strata-specific outcome rates where the weights represent the strata-specific sizes of the standard population. The indirect standardized rates were estimated as the strata-specific outcome rates from the standard population to derive expected outcome rates in the index population for the four different outcomes. For both rates 95% confidence intervals were calculated.[17-19]

We only presented direct standardized rates, unless indirect rates showed contrasting results.

RESULTS

Before the introduction of the Sophia birth centre, 443 (21%) of women planned a home delivery and 1391 (67%) planned a hospital delivery. After the introduction of the birth centre, 470 (12%) women planned a delivery at home, 1583 (41%) at the hospital, and 1671 (44%) at the birth centre.

After the introduction of the birth centre, women who planned birth at home were significantly more likely to be multiparous or had taken more often preconceptional folic

Table 5.1 Characteristics of women under supervision of a midwife (primary care) at the onset of labour, according to planned place of birth and the period before and after the introduction of the midwife-led birth centre

	Birth Centre Area Before introduction				Birth Centre Area After introduction							
	Home (n)	%	Hospital (n)	%	Home (n)	%	Hospital(n)	%	Birth Centre	%		
	443	21%	1,391	67%	470	12%	1,583	41%	1,671	44%		
Variable						A		B			C	D
Parity						*		*			*	*
Missing	1	0%	0	0%	1	0%	1	0%	6	0%		
Primiparous	246	56%	850	61%	191	41%	852	54%	970	58%		
Multiparous	196	44%	541	39%	278	59%	730	46%	695	42%		
Maternal Age								*			*	*
Missing	22	5%	73	5%	3	1%	17	1%	22	1%		
<19 years	9	2%	48	3%	4	1%	48	3%	55	3%		
20-25 years	28	6%	185	13%	24	5%	142	9%	205	12%		
25-34 years	279	63%	870	63%	306	65%	1,028	65%	1,055	63%		
>35 years	105	24%	215	15%	133	28%	348	22%	334	20%		
Ethnic background											*	*
Missing	42	9%	252	18%	54	11%	381	24%	218	13%		
Western	351	79%	723	52%	352	75%	783	49%	584	35%		
Non Western	50	11%	416	30%	64	14%	419	26%	869	52%		
Neighbourhood											*	*
Privileged neighbourhood	293	66%	802	58%	321	68%	963	61%	790	47%		
Underprivileged neighbourhood	150	34%	589	42%	149	32%	620	39%	881	53%		
Married status								*			*	*
Missing	8	2%	0	0%	5	1%	40	3%	28	2%		
Married/living together	398	90%	1,173	84%	435	93%	1,363	86%	1,303	78%		
Single	37	8%	218	16%	30	6%	180	11%	340	20%		
Education											*	
Missing	na		na		373	79%	1,348	85%	1,006	60%		
Low					7	1%	20	1%	52	3%		
Middle					17	4%	77	5%	276	17%		
High					73	16%	138	9%	337	20%		

BMI														*
Missing	11	2%	18	1%	18	4%		55	3%		39	2%		
<20	67	15%	205	15%	79	17%		230	15%		251	15%		
20-30	339	77%	1,048	75%	350	74%		1,165	74%		1,281	77%		
>30	26	6%	120	9%	23	5%		133	8%		100	6%		
Folic Acid use							*			*			*	*
Missing	17	4%	53	4%	15	3%		31	2%		43	3%		
No	76	17%	338	24%	46	10%		264	17%		303	18%		
Preconceptional	53	12%	177	13%	126	27%		349	22%		429	26%		
Only during pregnancy	297	67%	823	59%	283	60%		939	59%		896	54%		
Antenatal care > 14 weeks										*			*	*
No	359	81%	1,084	78%	411	87%		1,353	85%		1,357	81%		
Yes	84	19%	307	22%	59	13%		230	15%		314	19%		
Gestational Age										*				
Missing	0	0%	3	0%	0	0%		1	0%		0	0%		
<34wk	5	1%	10	1%	2	0%		17	1%		12	1%		
35-36wk	4	1%	19	1%	4	1%		32	2%		33	2%		
37wk	11	2%	61	4%	13	3%		67	4%		69	4%		
38-41 wk	422	95%	1,286	92%	451	96%		1,464	92%		1,553	93%		
>41 wk	1	0%	12	1%	0	0%		2	0%		4	0%		
Big3 outcomes														
SGA	29	7%	107	8%	23	5%		121	8%		120	7%		
Premature	6	1%	28	2%	6	1%		46	3%		44	3%		
Congenital abnomality	8	2%	59	4%	2	0%		7	0%		24	1%		
Combination Big3	3	1%	13	1%	1	0%		3	0%		8	0%		
Total Big3	46	10%	207	15%	32	7%		177	11%	*	196	12%	*	

*p < 0.05 (chi square test or Fisher exact test).
**Totals may not add up to 100 because of rounding error.
na = not available.
A = home before vs. after the introduction of the midwife-led birth centre.
B = hospital before vs. after the introduction of the midwife-led birth centre.
C = birth centre vs. home (after the introduction of the midwife-led birth centre).
D = birth centre vs. hospital (after the introduction of the midwife-led birth centre).

acid (all favourable characteristics) compared to this group before the introduction (see column A). After the birth centre introduction, also the group of women who planned birth at the hospital showed a more favourable profile compared to this group before. This group was significantly more often multiparous, older, married or living together, had taken preconceptional folic acid, and had received antenatal care before 14 weeks of gestation (see column B).

The women who planned birth at the birth centre showed more unfavourable characteristics compared to home as well as hospital birth in the same period (more likely to be nulliparous, of younger age, of non Western origin, from unprivileged neighbourhoods, single, did more often not take preconceptional folic acid, and more often received antenatal care after 14 weeks of gestation (see column C + D)).

In the birth centre women, neonatal case mix compared unfavourable too; the total prevalence of Big3 was higher (planned home 7% vs. planned hospital 11% vs. planned birth centre 12%).

After the introduction of the birth centre, the intrapartum and early neonatal mortality decreased combining all deliveries regardless of place (4/1834=2.2‰ vs. 2/3724=0.5‰).

Intrapartum related neonatal morbidities were significantly more common in the period before the introduction of the birth centre (91/1834=5.0% vs. 140/3724=3.8%; data not shown), as described in table 5.2. After the introduction significantly more intrapartum related neonatal morbidities occurred in planned birth centre births compared to planned home or planned hospital births (5.3% vs. 1.9% vs. 2.7% respectively). This difference was primarily related to the difference of neonatal infections (1.80% vs. 0.64% vs. 0.57% respectively)

Intrapartum related maternal morbidities were also lower in the period after the introduction (153/1834=8.3% vs. 270/3724=7.3%). After the introduction of the birth centre they were highest among planned hospital births. Interventions were about equal in both periods (14.9% vs. 14.3%). In the period after the introduction, interventions occurred significantly more often in birth centre deliveries compared to planned home or planned hospital births (16% vs. 11% vs. 14% respectively). Although not significantly, caesarean delivery occurred more often in transferred patients from planned birth centre deliveries compared to transferred patients from planned home or planned hospital births (6% vs. 3% vs. 5% respectively).

Table 5.2 Outcomes of women under supervision of the midwife (primary care) at the onset of labour according to planned place of birth and the period before and after the introduction of the midwife-led birth centre

	Birth Centre Area Before introduction				Birth Centre Area After introduction							
	Home (n)	%	Hospital (n)	%	Home (n)	%	Hospital(n)	%	Birth Centre	%		
	443	21%	1,391	67%	470	12%	1,583	41%	1,671	44%		
Variable						A		B			C	D
Intrapartum related neonatal morbidities												
Neonatal encephalopathy	0	0.00%	0	0.00%	0	0.00%	0	0.00%	1	0.06%		
Brachius plexus injury	0	0.00%	0	0.00%	0	0.00%	1	0.06%	0	0.00%		
Fractured clavicle	0	0.00%	4	0.29%	0	0.00%	2	0.13%	4	0.24%		
Cephalohaematoma	0	0.00%	0	0.00%	0	0.00%	0	0.00%	0	0.00%		
Neonatal infection	3	0.68%	17	1.22%	3	0.64%	9	0.57%	30	1.80%		
Low Apgar score (<7 after 5 minutes)	4	0.90%	24	1.73%	4	0.85%	20	1.26%	29	1.74%		
Neonatal unit admission	10	2.26%	48	3.45%	6	1.28%	17	1.07%	35	2.09%		
Other trauma related to birth	0	0.00%	4	0.29%	0	0.00%	0	0.00%	3	0.18%		
Total intrapartum related neonatal morbidities	13	2.93%	78	5.61%	9	1.91%	43	2.72%	88	5.27%		
Intra partum & early neonatal death (24hrs)												
No	441	100%	1,389	100%	470	100%	1,582	100%	1,670	100%		
Yes	2	0.45%	2	0.14%	0	0.00%	1	0.06%	1	0.06%		
Intrapartum related maternal morbidities												
Postpartum haemorrhage >1,000cc	24	5%	80	6%	15	3%	86	5%	82	5%		
Third/fourth degree rupture	16	4%	33	2%	9	2%	35	2%	43	3%		
Total intrapartum related maternal morbidities	37	8%	107	8%	22	5%	118	7%	123	7%		
Interventions												
None	395	89%	1,166	84%	420	89%	1,368	86%	1,403	84%		
Vacuum or Forceps extraction	34	8%	159	11%	35	7%	134	8%	173	10%		
Caesarean section	14	3%	66	5%	15	3%	81	5%	95	6%		

*p < 0.05 (chi square test or Fisher exact test).
***Totals may not add up to 100 because of rounding error.
na = not available.
A = home before vs. after the introduction of the midwife-led birth centre.
B = hospital before vs. after the introduction of the midwife-led birth centre.
C = birth centre vs. home (after the introduction of the midwife-led birth centre).
D = birth centre vs. hospital (after the introduction of the midwife-led birth centre).

Figure 5.1 shows the birth weight distribution for all low risk singleton pregnant women, under supervision of a community midwife at the onset of labour, between 2000-2007 in Rotterdam, according to the planned place of birth. This was also shown for the periods before and after the introduction of the birth centre. In all periods, weights below the 10th percentile were lowest in planned home deliveries. The proportion of weights below the 2.3rd percentile decreased substantially in planned home deliveries after the introduction of the birth centre (2.9% (95% CI 1.7-5.0%) to 0.9%(0.3-2.2%)).

After the introduction of the birth centre, standardized intrapartum and early neonatal mortality rate decreased for all the different planned places of birth (taking the period before introduction as standard, see Table 5.3). Also intrapartum related neonatal morbidity and intrapartum related maternal morbidity rate decreased for all the different planned places of birth. Standardized intervention rates increased for home births and decreased for hospital births, and remained similar for planned birth centre births. However, standardized caesarean delivery rates increased for all the different planned places of birth. Direct and indirect standardization provided similar results.

When comparing the planned places of birth after the introduction of the birth centre, standardized intrapartum and early neonatal mortality rates and intrapartum related neonatal morbidity increased for birth centre births and decreased for home births

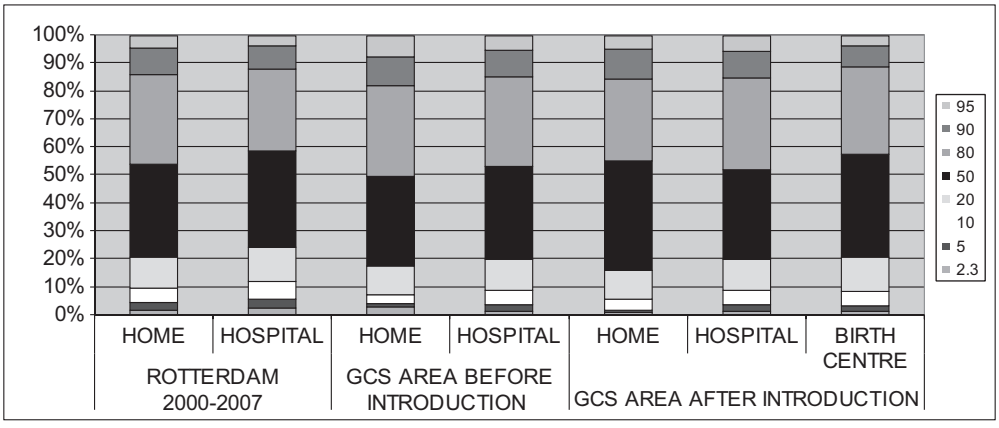


Figure 5.1 Birth weight distribution for all low risk singleton pregnant women, under supervision of a community midwife at the onset of labour, between 2000-2007 in Rotterdam, according to the planned place of birth.

Table 5.3 Mother and child outcomes according to planned place of birth (home, hospital and birth centre) after standardization for parity, age, etnicity and the presence of Small for Gestation Age (<p10)

	Birth Centre Area Before introduction		Birth Centre Area After introduction			Direct standardization								
	Home (n)	Hospital (n)	Home (n)	Hospital (n)	Birth Centre (n)	Home (vs. Home before introduction)			Hospital (vs. Hospital before introduction)			Birth Centre (vs. Hospital before introduction)		
	443	1,391	470	1,583	1,671	95%CI			95%CI			95%CI		
Outcome						RR	low	high	RR	low	high	RR	low	high
Intra partum & early neonatal death	2	2	0	1	1	0.00	0.00	0.00	0.31	0.31	0.31	0.86	0.86	0.86
Adverse child outcome birth related*	13	78	9	43	88	0.62	0.61	0.63	0.46	0.45	0.47	0.81	0.80	0.82
Adverse mother outcomes**	37	107	22	118	123	0.52	0.50	0.54	0.96	0.94	0.97	0.90	0.89	0.92
Interventions***	48	225	50	215	268	1.32	1.29	1.36	0.89	0.87	0.91	1.01	0.99	1.03
Caesarean delivery	14	66	15	81	95	1.37	1.35	1.38	1.19	1.17	1.20	1.14	1.13	1.15

* Adverse child outcome birth related is defined as: neonatal encephalopathy, brachius plexus injury, fractured clavicle, cephalohaematoma, neonatal infection, low Apgar score (<7 after 5 minutes), neonatal hospital admission, other trauma related to birth.
** Adverse mother outcome is defined as: post partum haemorrhage > 1,000cc or third or fourth degree rupture.
*** Interventions are defined as; receiving a medical intervention (vacuum extraction, forceps extraction, or caesarean delivery).

Table 5.4 Mother and child outcomes according to planned place of birth (home, hospital and birth centre) after direct standardization with the population planning a hospital birth as reference (standardization factors: parity, age, etnicity and the presence of Small for Gestation Age(SGA))

	Birth Centre Area After introduction			Direct standardization					
	Home (n)	Hospital (n)	Birth Centre (n)	Home vs. Hospital			Birth Centre vs. Hospital		
	470	1,583	1,671	95%CI			95%CI		
Outcome		(REF)		RR	low	high	RR	low	high
Intra partum & early neonatal death	0	1	1	0.00	0.00	0.00	2.23	0.27	4.19
Adverse child outcome birth related*	9	43	88	0.64	0.50	0.77	1.69	1.26	2.12
Adverse mother outcomes**	22	118	123	0.65	0.58	0.73	0.94	0.76	1.12
Interventions***	50	215	268	1.07	1.00	1.13	1.09	0.94	1.24
Caesarean Delivery	15	81	95	0.86	0.85	0.87	0.94	0.93	0.95

* Adverse child outcome birth related is defined as± neonatal encephalopathy, brachius plexus injury, fractured clavicle, cephalohaematoma, neonatal infection, low Apgar score (<7 after 5 minutes), neonatal hospital admission, other trauma related to birth.
** Adverse mother outcome is defined as: post partum haemorrhage > 1.000cc or third or fourth degree rupture.
*** Interventions are defined as; receiving a medical intervention (vacuum extraction, forceps extraction, or caesarean delivery).

compared to planned hospital (see Table 5.4). The intrapartum related maternal morbidity rate decreased for both home and birth centre births. Standardized intervention rates increased for home and birth centre births, but caesarean delivery decreased for planned home and birth centre births. Direct and indirect standardization provided similar results.

DISCUSSION

The introduction of a midwife-led (homelike) birth centre led to a redistribution of intended place of midwife-led births. Low risk women planning their delivery in the midwife-led birth centre have a higher risk profile compared to low risk women who planned their delivery at home or in the hospital, all under supervision of the midwife. This was confirmed by the birth weight distribution. Although the study did not have sufficient power to interpret the observed change in intrapartum and early neonatal death after the introduction of the birth centre (0.22% vs. 0.05%), the decreasing trend observed in planned home, hospital and birth centre births, suggests on average better care through more adequate selection. Alternative explanations of this trend should yet be explored (e.g. increased quality of care through the implementation of risk-led care, and the availability of acute obstetric and pediatric care). A similar trend was observed in intrapartum related neonatal morbidities, decreasing from 5.0% to 3.8%. Standardized rates showed the largest decrease in intended planned hospital births compared to planned home and birth centre births. A similar trend was observed in intrapartum related maternal morbidities (decreasing from 8.3% to 7.3%). Standardized rates showed the largest decrease in the planned birth centre births. The rate of interventions in our entire study population was 14.5%. Total intervention rates appeared unaffected by the introduction of the birth centre. Standardized rates of interventions were higher in planned home births, lower in planned hospital births and at the expected rate in birth centre births. Surprisingly, this trend was not observed for caesarean deliveries, which increased after the introduction of the birth centre.

Our cohort study showed some strengths. We used an intention-to-treat-like approach without ex post exclusion of unsuitable midwife cases to create a fair, unbiased comparison of planned home, hospital and birth centre births in this observational context. These unsuitable cases show poorer outcomes, and should therefore be included.[9]

Our case mix adjustment, using the presence of SGA, proved to be essential. The assumption of comparability across planned places of birth appeared not to be justified, judging from the unequal risk profiles, with home deliveries clearly representing the healthiest group. Self selection by the pregnant women can coincide with implicit or explicit selection by

the midwife who may tend to 'refer' to a hospital or birth centre if she feels uncomfortable with the risk level at home. Our adjustment of neonatal outcomes with birth weight was done before in similarly standardized analyses.[20]

Both direct and indirect standardization rates were applied to avoid the effect of accidental outlying subgroups. Direct standardization (weights taken from the index population) gives greater comparability but requires more data. Indirect standardization (weights taken from the standard population) requires fewer data but provides less comparability (unless the distribution of the standardization variable is identical across the study populations, in which case standardization is unnecessary since the crude mortality rates could have been compared directly).

As our results show the complete experience of the introduction of a midwife-led birth centre adjacent to the hospital on maternal and perinatal outcome in the north side of Rotterdam, its generelizability is mainly for urban areas in the Netherlands. Some study limitations merit discussion. A randomized controlled trail would be the superior design to address our research question. However when homebirth was part of a trial, participation hampered[21,22] and introduced selective participation which limits generalizability. Moreover if following one's choice impacts outcome, as expected here, estimates of setting effects are biased too.[21,23] Observational studies as ours are therefore indispensable, despite their shortcomings, in particular the difficulty to overcome the confounding by indication phenomenon.

Secondly, the study did not have sufficient power to assess intrapartum and early neonatal death. Thirdly, besides the introduction of the birth centre other reasons might be responsible for the improvements observed in the historical comparison. Fourthly, while the presence of SGA was used as additional case mix adjustment, our case mix adjustment could be further improved by detailed risk factors. We cannot rule out remnant confounding by indication as little is known on the factors underlying choice of setting. In addition, we categorized women into Dutch and non-Dutch women for reasons of power, where adverse outcomes may differ among the different ethnic groups (e.g. the increased prevalence of postpartum haemorrhage[24]). When the registration of detailed risk factors is improved one may consider a case-control model to compare the different places of birth.

Lastly, this study largely neglects the emotional aspects and the aspect of autonomy when comparing places of birth. The choice of the place of birth is largely upon the pregnant woman. The current growth of the share of birth centre and hospital births suggests overall positive balance of these effects, in particular since the economic incentive is in favour of home delivery. Studies assessing the mother's opinion show that any increase of medical safety easily outweighs other benefits, i.e. emotional aspects.[25]

Mixed effects of the introduction of the birth centre are observed since women planning their pregnancy in the midwife-led birth centre apparently have a higher risk profile compared to women who planned their pregnancy at home or in the hospital. A similar trend is observed in the Birthplace cohort[26], while most other international studies show the opposite.[27-31] Differential use of these options can be explained by several factors, either intentional or coincidental. After the introduction of the birth centre, low risk women could not plan their delivery in the hospital adjacent to the birth centre anymore, but are still able to plan their delivery in one of the other nearby hospitals. This may have led to a shift from the previously planned hospital births to the birth centre. Secondly, our birth centre aims to provide risk led care, with special attention to ethnic minorities and women with of low social economic background. This encourages caregivers to offer the higher risk women (among the low risk group) more explicitly the option of a birth centre delivery. Furthermore, in contrast to planned hospital births, women can receive postpartum care for at least four days in the midwife-led birth centre as an option. This may also attract caregivers of high risk groups.

The decreasing trend in mortality rate after the introduction of the birth centre (0.22% vs. 0.05%) should be interpreted with caution due to small numbers. If this, however, truly represents a decrease, it may in part be explained by the beneficial effect of local and national initiatives to lower perinatal mortality, in particular improved risk selection across all delivery options.[32-34] Our study showed an overall decreasing intrapartum related neonatal morbidity rate after the birth centre had been introduced (5.0% vs. 3.8%). Standardized rates showed a larger decrease of neonatal morbidity in planned hospital births compared to planned birth centre and home births. This may be due to residual confounding or an actual positive setting effect of planned hospital births, which is also observed by Evers et al.[35] The observed decreasing trend suggests on average better care through more adequate risk selection. Alternative explanations are the increased quality of first and second trimester screening, a higher awareness of malpractice procedures of both midwifes and obstetricians, and the continuous presence of senior obstetricians in perinatal centres. For unknown reasons neonatal infections were highest for planned birth centre births compared to planned hospital and home births within the same period. Differences in local protocols, definitions and management should be further explored.

The total prevalence of intrapartum related maternal morbidities also decreased after the introduction of the birth centre (8.3% to 7.3%). The decreasing trend observed in planned home, hospital and birth centre midwife-led births, may suggest on average better care through more adequate selection of women and/or more adequate management in all settings, e.g. exact measurement of blood loss and early use of intramuscular oxytocin.

Overall intervention rates were not affected by the introduction of the birth centre. While the underlying pattern suggests better fit of risk profile to setting and a better fit of risk profile to the choice for interventions, the introduction of an extra hospital-based facility did not represent an up- or downward pressure towards intervention rates in general. Previous studies on birth centres showed lower intervention rates combined with an equal, or even better, performance.[28-30] These studies, however, did not or only partially adjust for case mix differences.[28-30] The few available randomized controlled comparisons also showed lower rates[12,13], or at most an equal intervention rate with equal perinatal outcomes.[27,36,37] As these trials suffered from non participation or study small size[12,13,27,36,37], and showed difficult to combine results[11] , our study adds observational evidence from a large unselected cohort.

CONCLUSIONS

The introduction of a midwife-led birth centre led to a redistribution of women planning their midwife-led delivery at home, at the hospital, or at the local birth centre. Women opting for a delivery in the midwife-led birth centre had the most unfavourable risk profile.

Intrapartum and early neonatal mortality and intrapartum related neonatal and maternal morbidities tended to decrease, while overall intervention rates were unaffected. The introduction of the midwife-led birth centre seems to benefit the outcome of midwife-led deliveries. This change could be interpreted by the redistribution of the higher risk women among the low risk population intending birth at the birth centre instead of home. Alternative explanations of this trend should yet be explored. We therefore conclude that a risk profile should guide the choice of the proper place of delivery. The innovation of a midwife-led birth centre might be a first step to shared care where both midwives and obstetricians work together and learn from each other's expertise. This study may inform policy makers how to organize delivery care for low risk women.

REFERENCES

1. CVZ. CV. List of Obstetric Indications (in Dutch: Verloskundig Vademecum) 2003. Diemen (http://www.knov.nl/docs/uploads/Verloskundig_Vademecum_2003.pdf). 2003.
2. Nederland SPR. Perintale Zorg in Nederland 2008. Utrecht: 2011.
3. Ackermann-Liebrich U, Voegeli T, Gunter-Witt K, et al. Home versus hospital deliveries: follow up study of matched pairs for procedures and outcome. Zurich Study Team. Bmj. 1996 Nov 23;313(7068):1313-8.

4. Hutton EK, Reitsma AH, Kaufman K. Outcomes associated with planned home and planned hospital births in low-risk women attended by midwives in Ontario, Canada, 2003-2006: a retrospective cohort study. Birth. 2009 Sep;36(3):180-9.

5. Janssen PA, Lee SK, Ryan EM, et al. Outcomes of planned home births versus planned hospital births after regulation of midwifery in British Columbia. Cmaj. 2002 Feb 5;166(3):315-23.

6. Lindgren HE, Radestad IJ, Christensson K, Hildingsson IM. Outcome of planned home births compared to hospital births in Sweden between 1992 and 2004. A population-based register study. Acta Obstet Gynecol Scand. 2008;87(7):751-9.

7. Olsen O. Meta-analysis of the safety of home birth. Birth. 1997 Mar;24(1):4-13; discussion 4-6.

8. Wiegers TA, Keirse MJ, van der Zee J, Berghs GA. Outcome of planned home and planned hospital births in low risk pregnancies: prospective study in midwifery practices in The Netherlands. Bmj. 1996 Nov 23;313(7068):1309-13.

9. van der Kooy J, Poeran J, de Graaf JP, et al. Planned home compared with planned hospital births in the Netherlands: intrapartum and early neonatal death in low-risk pregnancies. Obstet Gynecol. 2011 Nov;118(5):1037-46.

10. Chervenak FA, McCullough LB, Brent RL, Levene MI, Arabin B. Planned home birth: the professional responsibility response. Am J Obstet Gynecol. Jan;208(1):31-8.

11. Hodnett ED, Downe S., Edwards N., Walsh D. Home-like versus conventional institutional settings for birth. The Cochrane Collaboration. 2009.

12. Hundley VA, Cruickshank FM, Lang GD, et al. Midwife managed delivery unit: a randomised controlled comparison with consultant led care. Bmj. 1994 Nov 26;309(6966):1400-4.

13. Waldenstrom U, Nilsson CA, Winbladh B. The Stockholm birth centre trial: maternal and infant outcome. Br J Obstet Gynaecol. 1997 Apr;104(4):410-8.

14. Coyle KL, Hauck Y, Percival P, Kristjanson LJ. Normality and collaboration: mothers' perceptions of birth centre versus hospital care. Midwifery. 2001 Sep;17(3):182-93.

15. Visser GH, Eilers PH, Elferink-Stinkens PM, Merkus HM, Wit JM. New Dutch reference curves for birthweight by gestational age. Early Hum Dev. 2009 Dec;85(12):737-44.

16. Bonsel GJ, Birnie E, Denktas S, Poeran J, Steegers EAP. Dutch report: Lijnen in de Perinatale Sterfte, Signalementstudie Zwangerschap en Geboorte 2010 (http://www.nvk.nl/Nieuws/Dossiers/DossierPerinatalezorg.aspx). Rotterdam: Erasmus MC; 2010.

17. Armitage P BG. Statistical Methods in Medical Research (3rd edition). Blackwell 1994.

18. Dobson AJ KK, Eberle E, Scherer J. Confidence intervals for weighted sums of Poisson parameters. Statistics in Medicine 1991;10:457-462.

19. Kahn HA SC. Statistical Methods in Epidemiology. 1989. New York: Oxford University Press.

20. Foster JE, Kleinman JC. Adjusting neonatal mortality rates for birth weight. Vital Health Stat 2. 1982 Aug(94):1-39.

21. Tyson H. Outcomes of 1001 midwife-attended home births in Toronto, 1983-1988. Birth. 1991 Mar;18(1):14-9.

22. Hendrix M, Van Horck M, Moreta D, et al. Why women do not accept randomisation for place of birth: feasibility of a RCT in The Netherlands. Bjog. 2009 Mar;116(4):537-42; discussion 42-4.

23. Dowswell T, Thornton JG, Hewison J, et al. Should there be a trial of home versus hospital delivery in the United Kingdom? Bmj. 1996 Mar 23;312(7033):753-7.

24. Bryant A, Mhyre JM, Leffert LR, Hoban RA, Yakoob MY, Bateman BT. The association of maternal race and ethnicity and the risk of postpartum hemorrhage. Anesth Analg. 2012 Nov;115(5):1127-36.

25. Bijlenga D, Birnie E, Bonsel GJ. Feasibility, Reliability, and Validity of Three Health-State Valuation Methods Using Multiple-Outcome Vignettes on Moderate-Risk Pregnancy at Term. Value Health. 2009 Mar 10.

26. Perinatal and maternal outcomes by planned place of birth for healthy women with low risk pregnancies: the Birthplace in England national prospective cohort study. BMJ. 2011;343:d7400.

27. Eide BI, Nilsen AB, Rasmussen S. Births in two different delivery units in the same clinic—a prospective study of healthy primiparous women. BMC Pregnancy Childbirth. 2009;9:25.

28. Gottvall K, Grunewald C, Waldenstrom U. Safety of birth centre care: perinatal mortality over a 10-year period. BJOG. 2004 Jan;111(1):71-8.

29. Gottvall K, Waldenstrom U, Tingstig C, Grunewald C. In-hospital birth center with the same medical guidelines as standard care: a comparative study of obstetric interventions and outcomes. Birth. 2011 Jun;38(2):120-8.

30. Laws PJ, Tracy SK, Sullivan EA. Perinatal outcomes of women intending to give birth in birth centers in Australia. Birth. 2010 Mar;37(1):28-36.

31. Waldenstrom U, Nilsson CA. Characteristics of women choosing birth center care. Acta Obstet Gynecol Scand. 1993 Apr;72(3):181-8.

32. Denktas S, Bonsel GJ, Steegers EA. [Perinatal health in Rotterdam, the Netherlands—experiences after 2 years of 'Ready for a baby']. Ned Tijdschr Geneeskd. 2012;156(29):A4289.

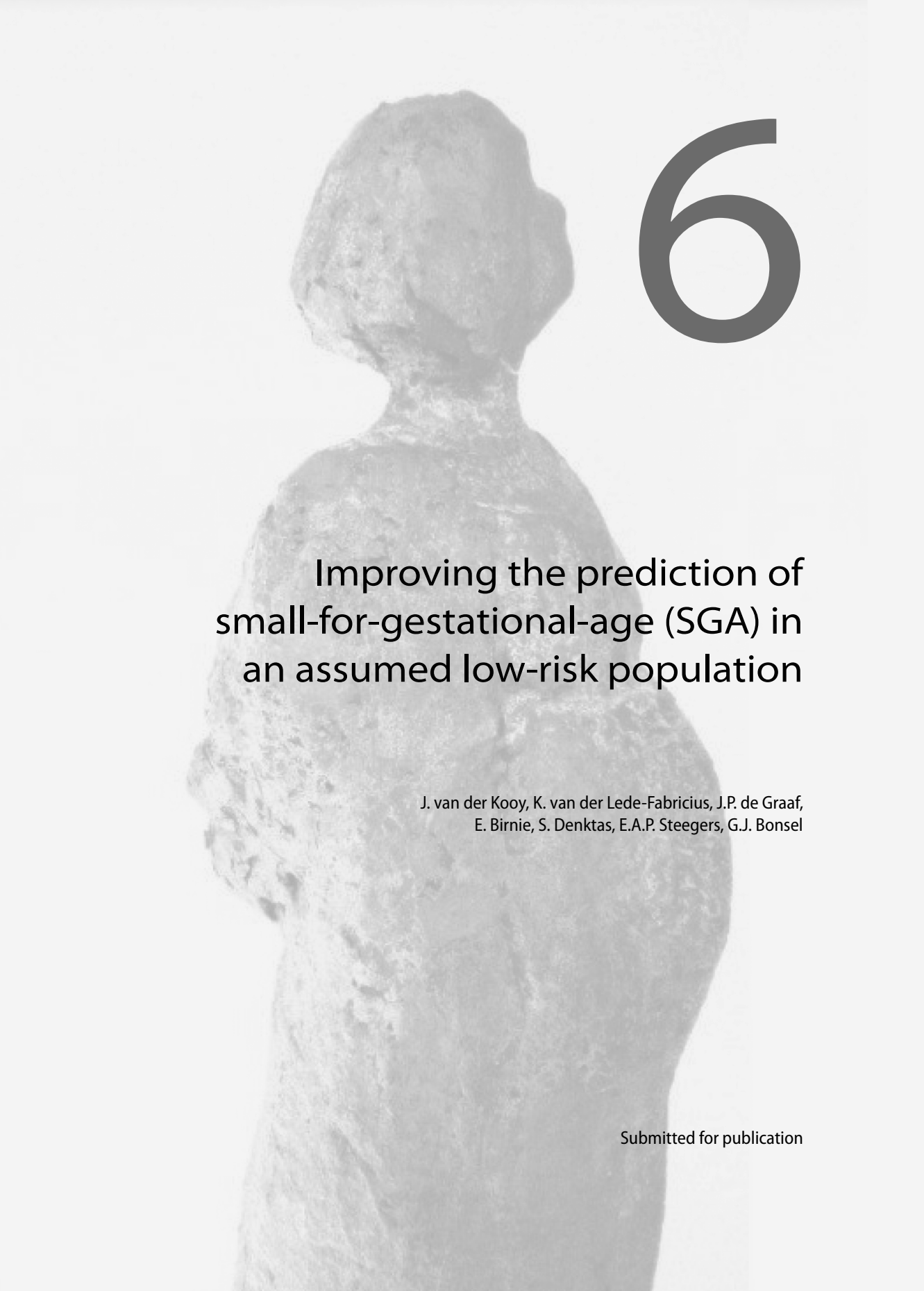
33. Denktas S, Bonsel GJ, Van der Weg EJ, et al. An urban perinatal health programme of strategies to improve perinatal health. Matern Child Health J. 2012 Nov;16(8):1553-8.

34. van der Velden J. Een goed begin. Veilige zorg rond zwangerschap en geboorte. http://www.knov.nl/docs/uploads/Advies_stg_zw_en_geb_Een_goed_beginpdf_spread-versie_voor_media.pdf, 2009.

35. Evers AC, Brouwers HA, Hukkelhoven CW, et al. Perinatal mortality and severe morbidity in low and high risk term pregnancies in the Netherlands: prospective cohort study. Bmj. 2010;341:c5639.

36. Begley C, Devane D, Clarke M, et al. Comparison of midwife-led and consultant-led care of healthy women at low risk of childbirth complications in the Republic of Ireland: a randomised trial. BMC Pregnancy Childbirth. 2011;11:85.

37. Bernitz S, Rolland R, Blix E, Jacobsen M, Sjoborg K, Oian P. Is the operative delivery rate in low-risk women dependent on the level of birth care? A randomised controlled trial. BJOG. 2011 Oct;118(11):1357-64.



6

Improving the prediction of small-for-gestational-age (SGA) in an assumed low-risk population

J. van der Kooy, K. van der Lede-Fabricius, J.P. de Graaf,
E. Birnie, S. Denktas, E.A.P. Steegers, G.J. Bonsel

Submitted for publication

ABSTRACT

Objective In the Netherlands perinatal mortality rates exceed the European average. The highest rates of perinatal mortality and morbidity are observed in the four largest cities, in particular in deprived neighbourhoods. The prevalence of small for gestational age (SGA) in assumed low-risk pregnancies at the onset of birth is still high (8%; which are 27% of all SGA cases) indicating that the performance of the current antenatal risk detection is suboptimal. We aimed to improve the detection of SGA by using the Rotterdam Reproductive Risk Reduction (R4U) as an adjuvant diagnostic tool in an assumed low-risk population.

Methods A prospective cohort study of all 1578 women who gave birth between October 2009 and December 2011 at the Sophia birth centre, Rotterdam, the Netherlands. SGA was defined as birth weight below the 10th percentile for gestational age, gender and parity. Multiple logistic regression analysis was used to develop a prediction model for undetected SGA, stratified for ethnic group. The performance of the prediction model was evaluated using six prognostic test characteristics.

Results The prevalence of undetected SGA at the onset of birth was 5.8% (95% CI 4.0-8.2) in Dutch and 10.4% (95% CI 8.4-12.6) in non-Dutch women. Factors predicting undetected SGA among Dutch women were: being unemployed (OR 3.16; 95% CI 1.29-7.74), living in a deprived neighbourhood (OR 2.78; 95% CI 1.29-6.01), having a history of prematurity (OR 7.37; 95% CI 1.46-37.29) and having a previous SGA (OR 5.67; 95% CI 1.92-16.73). Factors predicting undetected SGA in non-Dutch women were: being single (OR 2.22; 95% CI 1.38-3.56), case file social work (OR 3.78; 95% CI 1.40-10.22), nulliparity (OR 1.91; 95% CI 1.13-3.23) and having a previous SGA (OR 5.77; 95% CI 2.94-11.30). Diagnostic accuracy for the different thresholds increased mainly for Dutch women compared to the total group.

Conclusion The power to predict the risk of undetected SGA at birth in an assumed low-risk population seems optimal when both obstetric and social risk factors and ethnicity are taken into account. The R4U is a valuable adjuvant diagnostic tool for this purpose.

INTRODUCTION

In the Netherlands perinatal mortality rates have exceeded the European average for numerous years. In the four largest Dutch cities, i.e. Amsterdam, Rotterdam, The Hague and Utrecht, these rates are even higher, with the highest rates of perinatal mortality (18 per 1000 births) and morbidity being observed in deprived neighbourhoods.[1-5] In the Dutch system, independently operating community midwives provide care for assumed low-risk and medium-risk pregnant women. Assumed high-risk pregnant women are referred to the gynaecologist for evaluation and for antepartum and intrapartum care when the risk is confirmed.[6] The current selection of risk is primarily done by all midwives and based on the List of Obstetric Indications (LOI), a national indication list which mainly focuses on the presence of clinical (e.g. medical and obstetric) risk factors.[6] The purpose of risk selection is that the proportion of high risk pregnancies and deliveries under midwife responsibility is kept to a minimum. The proportion of small for gestational age (SGA) births under midwife's responsibility therefore serves as an indicator of the degree of suboptimal risk selection, since the presence of SGA is judged as a high risk pregnancy and should therefore be referred to the gynaecologist. Early detection and referral of SGA may improve perinatal outcome.[1,7-8] The rationale behind this study includes two reasons. Firstly, despite the current LOI-based antenatal risk selection, which mainly focuses on medical and obstetric risk factors, fundus height measurements, and the use of ultra sound when indicated, the prevalence of SGA in women who start birth under the supervision of a midwife is still about 8%, which are 27% of all SGA cases.[8] This suggests that the current process of risk selection should be improved. Secondly, several studies have shown that not only clinical risk factors but also pre-existent socioeconomic factors play a role in the occurrence of SGA. Timely recognition of these risks can be beneficial to detect SGA.[9-13] Other studies showed that predictors may differ for different ethnic groups.[14-16]

In order to improve the antenatal detection rates of SGA we developed an adjuvant diagnostic tool: the Rotterdam Reproductive Risk Reduction Questionnaire (R4U). The R4U is a checklist screening tool which recognizes both clinical and non-clinical (e.g. social, health care and lifestyle) factors. The R4U was developed within an Urban Perinatal Health Programme to Improve Perinatal Health.[17] The R4U can be used at several stages during pregnancy; the appropriate consecutive action depends on the stage and on the risk factors identified.

In Rotterdam, a midwife-led birth centre adjacent to a hospital was established in order to provide a safe environment for low-risk women to give birth. When women enter the birth centre, the R4U is used to detect a subgroup of women who could profit of specialised

care. One of the purposes of the R4U is to screen women who enter the birth centre on the presence of undetected SGA.

In this prospective study we evaluate if the detection of SGA in an assumed low-risk population can be improved by using the R4U as adjuvant diagnostic tool. We hypothesise that even at the onset of birth both clinical and non-clinical factors are of relevance in triage. In addition we assess whether the prediction of SGA is the same among different ethnic groups. As this was a proof of principle study no clinical interventions were implemented at this stage.

METHODS

Study design and setting

The study was a prospective cohort study. All women who started birth at the midwife led Sophia birth centre, Rotterdam, the Netherlands, between October 2009 and December 2011 were included. The centre was established in October 2009 adjacent to a hospital and is staffed by local community midwives to attend low risk pregnant women in a facilitated home-like environment. Women who start birth in the Sophia birth centre are low-risk assuming that the national LOI criteria have been applied timely and effectively. [6] A pregnancy is low-risk at the start of birth when it is a singleton pregnancy, ≥37.0 weeks gestational age, without a medical or obstetric history requiring referral to secondary or tertiary care prior to delivery.

Dutch law exempts institutional review of the Medical Ethics Committee since anonymized medical records were used retrospectively.

Data collection

Data were abstracted from a medical paper registration form, which was routinely filled in by midwives. The form contains data on maternal, child and process outcomes (e.g. interventions and referral) and risk factors from the Reproductive Risk Reduction Questionnaire (R4U). No personal data were abstracted (strict anonymous data).

R4U-B checklist

The R4U can be used antenatal, at the onset of birth and postnatal. The R4U used at the Sophia birth centre is called the R4U-B (birth). Most of the R4U-B is filled in at the birth centre at the start of birth, but some factors e.g. the pre-existing non-clinical factors can already be obtained in the antenatal stage.

The R4U-B consists of 60 variables grouped to five domains: social, health care, lifestyle, medical history and obstetric history. Of these 60 variables, 37 were selected for their known association with SGA based on literature and since these were routinely asked.[5,9-12,14,18-27] Of the 37 variables associated to SGA, 20 non-clinical variables were grouped to three domains. The social domain (11 variables) consists of: marital status, domestic violence, case file social work in the past 2 years, maternal employment status, work during first trimester, net household income (<€1000 or >€1000), paternal employment status, educational level, living in a deprived neighbourhood (based on 4-digit zip codes and a public list of deprived zip-code based neighbourhoods issued by the Dutch government. The list of deprived zip-code based neighbourhoods is based on 18 indicators. Those indicators refer to both (social and physical) problems as well as (social and physical) deprivation. Next to objective criterions (statistics and numbers), also subjective criterions (views of residents) have been used.[28]), ethnic background (Dutch/non-Dutch) and language proficiency. The health care domain (5 variables) consists of: health insurance (insured/uninsured), pregnancy planning, spontaneous or assisted conception, maternal age, and the gestational age at booking. The lifestyle domain (4 variables) consists of: smoking, alcohol use, recreational drug use, and pre-pregnancy body mass index (BMI; in kg/m2).

The remaining 17 clinical variables consisted of the medical domain (9 variables (e.g. history of medical disease, medication use during pregnancy)) and the obstetric domain (8 variables e.g. parity, history of operative vaginal delivery).

Definition of SGA

The presence of small for gestational age (SGA) was defined as birth weight below the 10th percentile for gestational age, gender and parity specific according to the national birth weight reference curves.[29] Birth weight reference curves were unavailable for specific ethnic groups, except for Surinamese-Hindustani newborns. Application of Hindustani-specific weight references for Surinamese-Hindustani children did not change our prediction models.[29] Hence, for parsimony we did not use ethnic specific percentiles.

Missing data

Of the 1578 women in our study population, 77 women were excluded because they had ≥20% missing values of their variables. Of the remaining 1501 women, 63 women were excluded because the outcome SGA was missing. The study population available for analysis consisted of 1438 women with 20 non-clinical and 17 clinical variables. Missing rates

per variable varied from 1% up to 29%. For our analysis missing variable data within the study population were replaced or imputed. For variables that were unrelated to other data, missing values were replaced by mean, median and mode, for numeric (normally distributed), ordinal and nominal values, respectively. Correlated variables ($p<0.05$) were imputed using multiple imputation. For each missing value five draws were performed providing five substituted data per missing variable, which in turn created five different imputed data sets. Analyses were performed separately on each imputed dataset and thereafter combined into one overall result. An association was considered significant only if the same significant association was observed across at least four of the imputed datasets. In that case we report on the median of the five coefficients with its 95% confidence interval (CI).[5,30]

Statistical analysis

The data were analysed with Statistical Package of Social Sciences version 20.0 for Windows (SPSS Inc, Chicago, IL, USA). Analyses were done for the total study group and for Dutch and non-Dutch women separately, because we expected different predictors for SGA in Dutch and non-Dutch women.[14-16]

First, we compared characteristics between Dutch and non-Dutch women using chi-square tests and Fisher Exact Tests. We calculated crude odds ratios (OR, 95% CI) for SGA for all variables separately (univariate logistic regression analysis). We then applied multivariable logistic regression analysis (forced entry method; and stepwise analyses with inclusion $p<0.05$ and exclusion $p>0.10$). All analyses were repeated with a stepwise forward and stepwise backward approach and finally inclusion of only predictive variables ($p<0.05$) for all imputed datasets. Results across the three approaches were similar unless stated otherwise. Regression coefficients were used to produce a woman's probability of having SGA at birth centre entry.

Finally, the performance of the prediction models was evaluated using the proportion of correctly predicted women. In order to assess the validity and the potential usefulness of this diagnostic tool, the following six diagnostic test characteristics were determined:

- Sensitivity, the proportion of SGA pregnancies correctly identified as such;
- Specificity, the proportion of non-SGA pregnancies correctly identified as such;
- Positive predictive value (PPV), the proportion of pregnancies correctly classified by the R4U-B as being a SGA pregnancy;
- Negative predictive value (NPV), the proportion of pregnancies correctly classified by the R4U-B as being a non-SGA pregnancy;

- Number needed to screen (NNS), the number of low-risk women needed to screen at birth centre entry to identify one case of SGA (1/absolute risk reduction);
- Diagnostic accuracy (DA), the proportion of presence or absence of SGA correctly classified as such by the R4U-B.

In addition we report the absolute number of women who were correctly identified as SGA (true positive), the number who were incorrectly identified as SGA (false positive), the number correctly identified as non-SGA (true negative) and the number who had SGA but were not detected (false negative).

We calculated test characteristics for the following thresholds: 0.10, 0.15, 0.20, 0.25 and 0.30 (e.g., a threshold of 0.20 means that all women with a predicted risk of having SGA ≥20% are considered to have SGA). We estimated the optimal thresholds by reviewing the test characteristics described above and considering the different clinical implications of false negatives and false positives; this judgement ultimately is subjective. [31-32] The performance of the various prediction modules of the total population was also tested using data-splitting, since no other population was available. In accordance with the principles of data-splitting, five datasets were distinguished by random sampling. Four datasets were used to fit the model and one dataset was used for validation. Each dataset was used both for validation and to fit the model.

RESULTS

Risk profiles

Table 6.1 describes the baseline characteristics of the study population by ethnic group. Of the 1,438 included women, 36% were Dutch, 59% non-Dutch and the remaining 5% was unknown. The observed prevalence of SGA births in the Sophia birth centre was 8.6% (95% CI 7.2-10.2%) and differed significantly between Dutch and non-Dutch women 5.8% (95% CI 4.0-8.2%) and 10.4% (95% CI 8.4-12.6%), respectively (p<0.001).

In all domains, the prevalences of risk factors differed between Dutch and non-Dutch women. Social risk factors were more common in non-Dutch women, except for work during the first trimester. Domestic violence and case file social work were not significantly different. Health care related risk factors were more common in non-Dutch women, except for assisted conception. Maternal age was not significantly different between the groups. The prevalence of lifestyle related risk factors smoking and recreational drug abuse was about the same for

Table 6.1 Characteristics and outcome of women in the total and stratified population

Variable	Total %	Dutch %	Non-Dutch %	p-value
Social domain				
Marital status (Single or not living together)	19.9	12.8	24.2	**
Missing	.9	1.0	.7	
Domestic violence (Yes)	4.1	3.9	4.5	NS
Missing	18.8	13.3	20.8	
Contact with social care (Yes)	2.2	1.9	2.4	NS
Missing	21.1	15.5	23.2	
Employment status (Unemployed)	15.9	8.3	20.2	**
Missing	1.3	.4	1.9	
Work during first trimester (Yes)	67.8	88.2	55.4	**
Missing	.6	.4	.8	
Net income (<1,000)	11.6	5.0	15.3	**
Missing	22.7	16.1	25.5	
Employment status partner (Unemployed)	10.9	5.2	14.4	**
Missing	4.0	1.2	5.4	
Educational level (Primary)	8.1	1.9	12.2	**
Missing	17.5	5.4	24.1	
Neighbourhood (Underprivileged neighbourhood)	50.2	35.4	59.1	**
Missing	0	0	0	
Etnic background (Dutch)	36.0			
Missing	4.9			
Language proficiency (Linguistic barrier)	9.0	0	14.8	**
Missing	3.6	0	5.4	
Health care domain				
Health insurance (No health insurance)	1.0	.0	1.8	*
Missing	7.3	6.8	7.6	
Planned pregnancy (Unplanned)	20.2	10.6	26.1	**
Missing	11.0	7.5	12.5	
Assisted conception (KID, IVF, OI, IUI, ICSI)	1.9	4.1	.6	**
Missing	3.8	3.7	3.8	
Maternal age (<18 years or >40 years)	95.7	94.4	96.5	NS
Missing	.1	0	0	
Gestational age at booking (≥ 14 weeks)	22.9	12.8	28.6	**
Missing	30.0	32.5	28.5	

Lifestyle domain				
Smoking (Before or throughout pregnancy)	20.1	19.7	20	NS
Missing	.9	1.2	.8	
Alcohol use (Before or throughout pregnancy)	12.1	22.4	6.4	**
Missing	1.3	1.7	.8	
Recreational drug use (Yes)	1.7	1.0	2.1	NS
Missing	1.9	2.3	1.6	
Pre-pregnancy BMI (<18 or >30)	28.2	19.1	32.8	**
Missing	4.3	4.3	4.4	
Medical domain				
Medical history (Yes)	23.4	26.3	21.3	*
Missing	.8	1.0	.7	
Surgery in history (Yes)	26.4	35.8	20.9	**
Missing	.6	.4	.7	
Blood transfusion (Yes)	1.7	1.7	1.9	NS
Missing	.7	.4	.7	
Preconception medication (Yes)	7.9	8.3	7.6	NS
Missing	1.6	2.1	1.4	
Medication during pregnancy (Yes)	15.5	12.6	17.4	*
Missing	1.5	2.3	1.1	
Preconception folic acid use (No)	53.5	33.8	65.3	**
Missing	3.9	3.7	4.1	
Chlamydia in history (Yes)	6.4	7.2	5.9	NS
Missing	5.2	4.1	5.5	
Chlamydia during pregnancy (Yes)	1	.6	.9	NS
Missing	8.1	7.4	8.4	
Psychiatric history (Yes)	11.4	14.7	10.0	*
Missing	1.8	1.9	1.4	
Obstetric history domain				
Parity (Primiparous)	54.0	65.4	47.1	**
Missing	0	0	0	
2 or more miscarriages in history (Yes)	4.3	4.1	4.1	NS
Missing	0	0	0	
Prematurity in history (Yes)	2.0	1.5	2.2	NS
Missing	6.3	5.6	6.9	
Low apgar in history (Yes)	.4	.2	.5	NS
Missing	17.9	12.4	20.9	

(*continued*)

Table 6.1 Continued

Variable	Total %	Dutch %	Non-Dutch %	p-value
SGA in history (Yes)	6.5	4.1	7.9	*
Missing	.1	0	.1	
Congenital abnormality in history (Yes)	2.4	1.9	2.8	NS
Missing	1.3	1.2	1.3	
Perinatal death/IUFD in history (Yes)	.4	.4	.5	NS
Missing	.1	.2	0	
History of VE/forceps (Yes)	6.7	7.5	6.2	NS
Missing	3.3	2.3	3.9	
Outcome				
SGA (Yes)	8.6	5.8	10.4	*

p-value Dutch vs. non-Dutch. * = p<0.05; ** = p<0.001; NS = not significant.
Determinants included (reference): Marital status (Married or living together), Domestic violence (No), Case file social work (No), Maternal employment status (Employed), Work during first trimester (No), Net income (>€1,000), Paternal employment status (Employed), Educational level (Secondary or higher), Neighbourhood (Privileged neighbourhood), Ethnic background (Dutch), Communication (Unimpeded), Health insurance (Yes), Pregnancy planning (Yes), Assisted conception (No), Maternal age (19-39 years), Gestational age at booking (<14 weeks), Smoking (No), Alcohol use (No), Recreational drug use (No), BMI (18-30), History of medical disease (No), History of surgery (No), Blood transfusion (No), Preconception medication use (No), Medication use during pregnancy (No), Preconception folic acid use (Yes), History of Chlamydia (No), Chlamydia during pregnancy (No), History of psychiatric disease (No), Parity (Multiparous), History of 2 or more miscarriages (No), History of History of prematurity (No), History of low Apgar score (No), History of SGA (No), History of congenital abnormality (No), History of perinatal death/IUFD (No), History of operative vaginal delivery (No).

both groups. Alcohol abuse was more common among Dutch women, while more non-Dutch women had a pre-pregnancy BMI of <18 or >30. Of the medical risk factors, medication use during pregnancy and non-use of preconception folic acid were more common among non-Dutch women. The risk factors history of medical disease, history of surgery and history of psychiatric disease were more common among Dutch women. In the obstetric domain, only parity and history of SGA were significantly different between both groups. More Dutch women were primiparous, while more non-Dutch women had a history of SGA.

Logistic regressions

Table 6.2 shows the results of univariate and multiple logistic regression analyses (entry method and stepwise forward approach). Results were different for the total and for the stratified populations, but all prediction models consisted of social and obstetric factors only. In the total population (Table 6.2a), three variables were significantly

Table 6.2a Association between individual risk factors and the incidence of SGA, results from five imputed datasets for the total population

	Total (R^2 = 9%)								
	Model 1			Model 2			Model 3		
	OR	95% CI	p	OR	95% CI	p	OR	95% CI	p
Social domain									
Marital status (Single or not living together)	2.86	1.94-4.21	**	2.21	1.31-3.75	*	2.78	1.85-4.16	**
Domestic violence (Yes)	1.71	0.79-3.69		1.13	0.47-2.72		nie		
Case file social work (Yes)	3.87	1.69-8.85	*	2.81	1.03-7.63	*	2.54	1.06-6.13	*
Maternal employment status (Unemployed)	2.22	1.45-3.39	**	1.32	0.60-2.91		nie		
Work during first trimester (Yes)	0.70	0.47-1.03		1.05	0.59-1.85		nie		
Net income (< €1,000)	1.53	0.98-2.41		0.70	0.39-1.27		nie		
Paternal employment status (Unemployed)	1.54	0.92-2.59		1.06	0.54-2.08		nie		
Educational level (Primary)	1.08	0.62-1.91		0.91	0.46-1.78		nie		
Neighbourhood (Underprivileged neighbourhood)	1.27	0.88-1.84		1.14	0.76-1.71		nie		
Ethnic background (Non-Dutch)	1.78	1.18-2.69	*	1.77	1.08-2.88	*	nie		
Language proficiency (Linguistic barrier)	0.87	0.43-1.75		0.86	0.38-1.96		nie		
Health care domain									
Health insurance (Uninsured)	1.64	0.37-7.35		1.38	0.27-7.05		nie		
Planned pregnancy (Unplanned)	1.93	1.29-2.89	*	1.14	0.68-1.91		nie		
Assisted conception (KID, IVF, OI, IUI, ICSI)	0.40	0.05-2.99		0.50	0.06-4.25		nie		
Maternal age (<18 years or >40 years)	1.64	0.76-3.53		1.26	0.54-2.93		nie		
Gestational age at booking (≥ 14 weeks)	0.93	0.60-1.46		0.70	0.43-1.15		nie		
Lifestyle domain									
Smoking (Before or throughout pregnancy)	1.74	1.16-2.62	*	1.20	0.75-1.93		nie		
Alcohol use (Before or throughout pregnancy)	1.33	0.80-2.23		1.57	0.88-2.81		nie		
Recreational drug use (Yes)	2.72	1.00-7.37	*	1.20	0.39-3.72		nie		
BMI (<18 or >30)	1.32	0.89-1.95		1.16	0.76-1.77		nie		

(*continued*)

Table 6.2a Continued

	Total (R^2 = 9%)								
	Model 1			Model 2			Model 3		
	OR	95% CI	p	OR	95% CI	p	OR	95% CI	p
Medical domain									
History of medical disease (Yes)	0.91	0.58-1.41		0.96	0.58-1.60		nie		
History of surgery (Yes)	0.80	0.52-1.24		1.00	0.62-1.61		nie		
Blood transfusion (Yes)	0.44	0.06-3.26		0.53	0.07-4.03		nie		
Preconception medication use (Yes)	1.16	0.60-2.22		1.16	0.53-2.54		nie		
Medication use during pregnancy (Yes)	1.12	0.69-1.84		1.13	0.64-1.99		nie		
Preconception folic acid use (Yes)	1.56	1.06-2.30	*	0.99	0.61-1.60		nie		
History of Chlamydia (Yes)	0.91	0.43-1.91		0.59	0.25-1.37		nie		
Chlamydia during pregnancy (Yes)	1.46	0.43-4.93		1.56	0.53-4.61		nie		
History of psychiatric disease (Yes)	0.99	0.55-1.77		0.81	0.42-1.56		nie		
Obstetric domain									
Parity (Primiparous)	1.04	0.72-1.50		1.75	1.05-2.90	*	nie		
History of 2 or more miscarriages (Yes)	1.37	0.61-3.08		1.35	0.55-3.32		nie		
History of prematurity (Yes)	2.26	0.85-6.03		2.56	0.83-7.93		nie		
History of low apgar score (Yes)	5.37	0.97-29.61		3.03	0.37-24.64		nie		
History of SGA (Yes)	4.32	2.65-7.06	**	5.63	3.04-10.40	**	4.54	2.74-7.52	**
History of congenital abnomality (Yes)	1.38	0.48-3.97		1.65	0.51-5.39		nie		
History of perinatal death/ IUFD (Yes)	2.13	0.25-18.36		1.29	0.13-13.16		nie		
History of VE or forceps (Yes)	0.56	0.23-1.42		0.74	0.25-2.13		nie		

Model 1: crude OR; Model 2: OR adjusted for social, health care, lifestyle, medical and obstetric risk factors (forced entry); Model 3: OR obtained with stepwise forward regression analysis.
nie = not in equation.
* = p<0.05; **= p<0.001.
Determinants included (reference): Marital status (Married or living together), Domestic violence (No), Case file social work (No), Maternal employment status (Employed).
Work during first trimester (No), Net income (>€1,000), Paternal employment status (Employed), Educational level (Secondary or higher), Neighbourhood (Privileged neighbourhood).
Ethnic background (Dutch), Communication (Unimpeded), Health insurance (Yes), Planned pregnancy (Yes), Assisted conception (No), Maternal age (19-39 years).
Gestational age at booking (<14 weeks), Smoking (No), Alcohol use (No), Recreational drug use (No), BMI (18-30), History of medical disease (No), History of surgery (No).
Blood transfusion (No), Preconception medication use (No), Medication use during pregnancy (No), Preconception folic acid use (Yes), History of Chlamydia (No), Chlamydia during pregnancy (No).
History of psychiatric disease (No), Parity (Multiparous), History of 2 or more miscarriages (No), History of prematurity (No), History of low Apgar score (No), History of SGA (No).
History of congenital abnormality (No), History of perinatal death/IUFD (No), History of operative vaginal delivery (No).

Table 6.2b Association between individual risk factors and the incidence of SGA, results from five imputed datasets for both Dutch and non-Dutch women

	Dutch (R^2 = 14%)									Non-Dutch (R^2 = 10%)								
	Model 1			Model 2			Model 3			Model 1			Model 2			Model 3		
	OR	95% CI	p	OR	95% CI	p	OR	95% CI	p	OR	95% CI	p	OR	95% CI	p	OR	95% CI	p
Social domain																		
Marital status (Single or not living together)	3.37	1.52-7.45	*	1.44	0.39-5.36		nie			2.48	1.59-3.88	**	2.20	1.19-4.09	*	2.22	1.38-3.56	**
Domestic violence (Yes)	0.87	0.11-6.69		0.82	0.07-9.06		nie			1.96	0.84-4.57		1.30	0.49-3.44		nie		
Case file social work (Yes)	1.87	0.23-15.22		0.16	0.00-5.10		nie			4.58	1.80-11.65	*	3.67	1.17-11.44	*	3.78	1.40-10.22	*
Maternal employment status (Unemployed)	4.23	1.77-10.07	*	4.33	0.49-38.09		3.16	1.29-7.74	*	1.66	1.02-2.69	*	1.05	0.43-2.56		nie		
Work during first trimester (Yes)	0.35	0.15-0.82	*	0.81	0.14-4.68		nie			0.99	0.63-1.54		1.01	0.55-1.86		nie		
Net income (< €1.000)	1.50	0.44-5.20		0.39	0.07-2.33		nie			1.32	0.80-2.16		0.82	0.42-1.61		nie		
Paternal employment status (Unemployed)	3.65	1.30-10.10	*	1.09	0.19-6.37		nie			1.09	0.60-2.00		0.92	0.41-2.07		nie		
Educational level (Primary)	0.00	0.00-.		0.00	0.00-.		nie			0.88	0.48-1.60		0.97	0.48-1.97		nie		
Neighbourhood (Underprivileged neighbourhood)	2.82	1.36-5.85	*	3.07	1.31-7.22	**	2.78	1.29-6.01	*	0.80	0.52-1.24		0.84	0.52-1.34		nie		
Ethnic background (Non-Dutch)																		
Language proficiency (Linguistic barrier)	nie			nie			nie			0.69	0.34-1.41		0.90	0.39-2.06		nie		

(*continued*)

Table 6.2b Continued

	Dutch (R^2 = 14%)									Non-Dutch (R^2 = 10%)								
	Model 1			Model 2			Model 3			Model 1			Model 2			Model 3		
	OR	95% CI	p	OR	95% CI	p	OR	95% CI	p	OR	95% CI	p	OR	95% CI	p	OR	95% CI	p
Health care domain																		
Health insurance (Uninsured)	nie			nie			nie			1.36	0.30-6.12		1.17	0.20-6.75		nie		
Planned pregnancy (Unplanned)	2.59	1.07-6.30	*	1.35	0.36-5.10		nie			1.58	1.00-2.50	*	1.15	0.64-2.07		nie		
Assisted conception (KID, IVF, OI, IUI, ICSI)	0.76	0.10-5.84		1.23	0.13-11.99		nie			0.00	0.00-.		0.00	0.00-.		nie		
Maternal age (<18 years or >40 years)	1.86	0.53-6.48		1.65	0.34-7.96		nie			1.73	0.65-4.62		1.16	0.39-3.44		nie		
Gestational age at booking (≥ 14 weeks)	0.70	0.21-2.36		0.29	0.06-1.48		nie			0.84	0.52-1.38		0.74	0.43-1.29		nie		
Lifestyle domain																		
Smoking (Before or throughout pregnancy)	2.19	1.02-4.69	*	1.42	0.51-3.97		nie			1.63	1.00-2.65		1.16	0.65-2.07		nie		
Alcohol use (Before or throughout pregnancy)	1.50	0.67-3.35		1.96	0.73-5.23		nie			1.93	0.94-3.97		1.33	0.58-3.08		nie		
Recreational drug use (Yes)	0.00	0.00-.		0.00	0.00-.		nie			3.48	1.21-10.00	*	1.80	0.50-6.44		nie		
BMI (<18 or >30)	1.03	0.43-2.43		0.53	0.17-1.60		nie			1.26	0.81-1.97		1.36	0.83-2.21		nie		

Medical domain																	
History of medical disease (Yes)	0.80	0.34-1.90		0.82	0.29-2.38		nie			1.00	0.59-1.69		0.95	0.51-1.76		nie	
History of surgery (Yes)	0.68	0.31-1.49		0.71	0.28-1.81		nie			0.98	0.58-1.68		1.17	0.66-2.08		nie	
Blood transfusion (Yes)	1.98	0.24-16.32		1.29	0.10-15.95		nie			0.00	0.00-.		0.00	0.00-.		nie	
Preconception medication use (Yes)	0.73	0.17-3.14		0.75	0.13-4.21		nie			1.35	0.65-2.82		1.60	0.64-3.97		nie	
Medication use during pregnancy (Yes)	1.56	0.62-3.93		2.03	0.65-6.39		nie			0.95	0.53-1.70		0.93	0.47-1.85		nie	
Preconception folic acid use (Yes)	2.04	1.00-4.13	*	1.45	0.55-3.85		nie			1.13	0.70-1.82		0.91	0.51-1.60		nie	
History of Chlamydia (Yes)	0.37	0.05-2.81		0.31	0.04-2.76		nie			1.24	0.54-2.81		0.71	0.29-1.77		nie	
Chlamydia during pregnancy (Yes)	0.00	0.00-.		0.00	0.00-.		nie			2.01	0.74-5.45		2.17	0.69-6.83		nie	
History of psychiatric disease (Yes)	1.39	0.55-3.48		1.12	0.34-3.72		nie			0.88	0.41-1.88		0.64	0.27-1.51		nie	
Obstetric domain																	
Parity (Primiparous)	0.80	0.38-1.65		1.37	0.40-4.72		nie			1.29	0.84-2.00		1.80	1.01-3.21	*	1.91	1.13-3.23 *
History of 2 or more miscarriages (Yes)	2.42	0.68-8.60		4.30	0.86-21.42					1.03	0.36-2.96		1.20	0.37-3.85		nie	
History of prematurity (Yes)	8.47	2.02-35.52	*	8.99	1.03-78.06	*	7.37	1.46-37.29	*	0.98	0.22-4.28		1.23	0.23-6.55		nie	

(*continued*)

Table 6.2b Continued

	Dutch (R^2 = 14%)									Non-Dutch (R^2 = 10%)								
	Model 1			Model 2			Model 3			Model 1			Model 2			Model 3		
	OR	95% CI	p	OR	95% CI	p	OR	95% CI	p	OR	95% CI	p	OR	95% CI	p	OR	95% CI	p
History of low apgar score (Yes)	0.00	0.00-.		0.00	0.00-.		nie			8.78	1.22-63.08	*	7.78	0.64-94.73		nie		
History of SGA (Yes)	6.50	2.37-17.80	**	12.05	2.54-57.15	*	5.67	1.92-16.73	*	3.54	2.02-6.21	**	5.17	2.56-10.45	**	5.77	2.94-11.30	**
History of congenital abnomality (Yes)	1.68	0.21-13.53		1.31	0.07-23.64		nie			1.24	0.36-4.23		2.05	0.52-8.04		nie		
History of perinatal death/IUFD (Yes)	16.06	0.98-262.75		4.52	0.10-203.06		nie			0.00	0.00-.		0.00	0.00-.		nie		
History of VE or forceps (Yes)	0.83	0.19-3.62		1.04	0.15-6.98		nie			0.48	0.15-1.55		0.47	0.11-2.01		nie		

Model 1: crude OR; Model 2: OR adjusted for social, health care, lifestyle, medical and obstetric risk factors (forced entry); Model 3: OR obtained with stepwise forward regression analysis.
nie = not in equation.
* = p<0.05; **= p<0.001.
Determinants included (reference): Marital status (Married or living together), Domestic violence (No), Case file social work (No), Maternal employment status (Employed), Work during first trimester (No), Net income (>€1,000), Paternal employment status (Employed), Educational level (Secondary or higher), Neighbourhood (Privileged neighbourhood), Ethnic background (Dutch), Communication (Unimpeded), Health insurance (Yes), Planned pregnancy (Yes), Assisted conception (No), Maternal age (19-39 years), Gestational age at booking (<14 weeks), Smoking (No), Alcohol use (No), Recreational drug use (No), BMI (18-30), History of medical disease (No), History of surgery (No), Blood transfusion (No), Preconception medication use (No), Medication use during pregnancy (No), Preconception folic acid use (Yes), History of Chlamydia (No), Chlamydia during pregnancy (No), History of psychiatric disease (No), Parity (Multiparous), History of 2 or more miscarriages (No), History of prematurity (No), History of low Apgar score (No), History of SGA (No), History of congenital abnormality (No), History of perinatal death/IUFD (No), History of operative vaginal delivery (No).

associated with an increased risk of SGA: single/not living together (OR 2.78; 95% CI 1.85-4.16), case file social work (OR 2.54; 95% CI 1.06-6.13) and history of SGA (OR 5.86; 95% CI 3.26-10.55).

When stratifying for ethnic group, both models consisted of four risk factors with a better performance compared to the unstratified population. Only history of SGA was universally present, all other factors differed among the stratified groups. In Dutch women, being unemployed (OR 3.16; 95% CI 1.29-7.74), living in a deprived neighbourhood (OR 2.78; 95% CI 1.29-6.01), history of prematurity (OR 7.37; 95% CI 1.46-37.29) and history of SGA (OR 5.67; 95% CI 1.92-16.73) were associated with SGA. In non-Dutch women, SGA was associated with single/not living together (OR 2.22; 95% CI 1.38-3.56), case file social work (OR 3.78; 95% CI 1.40-10.22), nulliparity (OR 1.91; 95% CI 1.13-3.23) and history of SGA (OR 5.77; 95% CI 2.94-11.30).

Predictive performance

Tables 6.3, 6.4 and 6.5 describe the test characteristics for different thresholds of predicted risk of having SGA for respectively Dutch, non-Dutch women and the total population. At a threshold of 0.10 for Dutch women, the specificity is 91% with a sensitivity of 38% and the NNS is 45. Of the 545 Dutch women, 20 women with SGA are missed and 46 are incorrectly identified as SGA. A threshold of 0.20 has a higher specificity (98%), but with a great loss of sensitivity (13%) and an increase of NNS (136). Only 10 women are incorrectly identified as SGA, and another 28 women with SGA are missed. For non-Dutch women, a threshold of 0.10 has a specificity of 78%, with a sensitivity of 50% and a NNS of 19. Of the 899 non-Dutch women only 46 women with SGA are missed, but 176 are incorrectly identified as SGA. A threshold of 0.20 has a higher specificity (92%), with a sensitivity of 29% and a NNS of 33. Only 66 women are incorrectly identified as SGA and 65 women with SGA are missed.

The prior probability of undetected SGA is higher among non-Dutch women, leading to different thresholds for both groups. We considered a predicted risk of SGA of 0.10 and 0.20 the optimal thresholds for Dutch and non-Dutch women respectively for the purpose of triage in this specific population. The predicted value of the total population using data- splitting showed an area under the curve ranging from 0.57-0.68.

Table 6.3 Test characteristics for Dutch women at different thresholds of predicted risk of having SGA

Threshold	Sensitivity	Specificity	PPV	NPV	NNS	DA	True positive	False positive	True negative	False negative
0.10	38%	91%	21%	96%	45	88%	12	46	467	20
0.15	28%	94%	23%	95%	61	90%	9	31	482	23
0.20	13%	98%	29%	95%	136	93%	4	10	503	28
0.25	13%	98%	29%	95%	136	93%	4	10	503	28
0.30	13%	100%	80%	95%	136	95%	4	1	512	28

PPV=positive predictive value; NPV=negative predictive value; NNS=numbers needed to screen; DA=diagnostic accuracy.

Table 6.4 Test characteristics for non-Dutch women at different thresholds of predicted risk of having SGA

Threshold	Sensitivity	Specificity	PPV	NPV	NNS	DA	True positive	False positive	True negative	False negative
0.10	50%	78%	21%	93%	19	75%	46	176	625	46
0.15	50%	78%	21%	93%	19	75%	46	176	625	46
0.20	29%	92%	29%	92%	33	85%	27	66	735	65
0.25	13%	98%	39%	91%	74	89%	12	19	782	80
0.30	10%	98%	38%	90%	99	89%	9	15	786	83

PPV=positive predictive value; NPV=negative predictive value; NNS=numbers needed to screen; DA=diagnostic accuracy.

Table 6.5 Test characteristics for the total population at different thresholds of predicted risk of having SGA

Threshold	Sensitivity	Specificity	PPV	NPV	NNS	DA	True positive	False positive	True negative	False negative
0.10	54%	76%	18%	95%	21	74%	67	312	1,002	57
0.15	27%	93%	28%	93%	44	88%	33	86	1,228	91
0.20	27%	93%	28%	93%	44	88%	33	86	1,228	91
0.25	11%	98%	37%	92%	103	91%	14	24	1,290	110
0.30	6%	99%	33%	92%	205	91%	7	14	1,300	117

PPV=positive predictive value; NPV=negative predictive value; NNS=numbers needed to screen; DA=diagnostic accuracy.

DISCUSSION

Main conclusion

The R4U as a non-invasive easy to administer adjuvant diagnostic tool improves the currently suboptimal detection of SGA in an assumed low-risk population at birth. The improved prediction of undetected SGA at the onset of birth in a midwife-led setting included only social and obstetric risk factors. This confirms our hypothesis that clinical factors alone are insufficient for optimal risk selection. We found different SGA predictors in Dutch and non-Dutch women, with a history of SGA as the single common factor. Consequently SGA prediction should be stratified for ethnic groups.

Strengths

A strength of this study is that we used an unselected prospective cohort. In addition, we assessed a broad spectrum of risk factors on a routine base, including both clinical and non-clinical factors. We studied the potential role of ethnicity as predictor, or alternatively using it as a stratifier or an adjustment factor when defining SGA percentiles. Stratification for ethnicity appeared optimal, as its effect went beyond a simple additive risk. This finding may contribute to improved guidelines for risk assessment, considering that current guidelines either assume no specific role for ethnicity at all, or a simple additive role.

Limitations

Some limitations merit discussion. The use of medical registration data implied substantial loss of variables in analysis due to missing data. The missing rate is particularly high for variables which combine 'inconvenience' if asked for, and which are prone for low awareness of their relevance. This was particular the case for non-clinical variables. The exclusion of women with incomplete data most likely yielded a relatively healthier sample for analysis, which in turn decreased additive predicting power. Our results, however, at least suggest that even routine practice data may suffice to improve detection.

A second limitation is the generalisability of our study, as our study included an urban population who started birth in a birth centre. The urban population differs from a rural population, since ethnic minorities and disadvantaged women are overrepresented in urban areas. This different distribution of risk factors not necessarily affects the selection of variables, but may alter the predictive power of these variables. The selection of women

at the time of birth was by design and represents a strength, since even at that stage risk differentiation is still feasible and effective. Thirdly, we only used Dutch weight references to define SGA, while weight references were available for Dutch and Hindustani children. Studies showed that Surinamese and Antillean newborns are smaller compared to Dutch newborns after adjusting for risk factors.[18,33-34] Using the Hindustani weight references for Surinamese-Hindustani children did not alter our prediction models. Furthermore, the predictive value of negative consequences of SGA of the 10th percentile of their own reference curves is still unknown. Data on the impact of ethnic differences in birth weight on mortality rates is conflicting.[35-36] We expect the use of Dutch weight references to be of little influence on our general study result, since a gradual effect exists between predictors and birth weight.

Fourthly, at this stage we only stratified for Dutch and non-Dutch women. More evidence on the need for specific birth weight reference curves will decide the need for larger datasets to establish further division of ethnic groups.

Finally, we showed that the detection of SGA improved by using the R4U. However, our study does not show whether improved detection actually leads to better outcomes since this was a proof-of-principle study.

Comparison with literature

While the issue of undetected SGA at birth gains attention with the emerging discussion of the clinical utility of third trimester ultrasound and other surveillance tools, to our knowledge no literature exists on undetected SGA prediction in an assumed low risk population.[37-39] Moreover, the independent SGA risk enhancement of non-clinical factors throughout pregnancy has gained little attention. We are aware of two SGA prediction studies that both reported social factors, such as ethnicity, marital status and maternal education to be of importance.[40-41] Clinical guidelines on antenatal care, such as the NICE guidelines, give due credit to the importance of such factors, but do not connect their presence with specific risks like SGA. These guidelines do, however, recommend intensified care for women with complex social factors.[42]

Unlike previous prediction models, risk factors in the lifestyle domain did not prove to have a strong impact at the onset of birth since smoking and abuse of recreational drugs was non-significant after adjustment for other risk factors. Possible explanations include the current extra antenatal surveillance in women with unhealthy lifestyle factors (e.g. smoking), resulting in referral to secondary care before the onset of birth. Also the social risk factors

may in part incorporate an unhealthy lifestyle and therefore social risk factors may reflect the association between an unhealthy lifestyle and SGA (mediation effect).[9,25-26]

Surprisingly, while the prevalence of premature birth in their history is equal among Dutch and non-Dutch women, a history of prematurity is a predictive factor of SGA only in Dutch women. It may be hypothesised that the underlying pattern of risks that leads to prematurity is the same as the underlying pattern leading to SGA in Dutch, but not in non-Dutch women where e.g. infectious causes are much more prominent. [43]

Optimal thresholds

As the R4U is a new diagnostic tool, there are no established thresholds for distinguishing high-risk from low-risk women. Any threshold implicitly assumes a balance between overtreatment (false positives) and undertreatment (false negatives) at that particular point. Here overtreatment implies referral of women, in the absence of SGA, to secondary care for intensified monitoring. This may include a confirmatory third trimester ultrasound. However, birth weight prediction in the third trimester is still challenging since estimation errors cannot be excluded.[44-46] If the absence of SGA is confirmed by the ultrasound in these initial false positives, they may still receive primary care. If the absence of SGA is not recognized by the ultrasound, these women may receive SGA-specialised care, which at this stage predominantly is focussed on a lower tolerance for any sign of deteriorating fetal health and could possibly lead to an increase in (unnecessary) interventions.[47] Undertreatment (false negatives) would imply unintentional denying access to SGA-specialised care, which at this stage, takes within the risk that SGA related risks (in particular fetal distress) are less easily detected since monitoring equipment (like cardiotocography, fetal scalp blood sample) are unavailable to the midwife. The major effect of SGA is not in the weight per se but in the exponential increase of perinatal mortality risk if other adverse conditions emerge.[1]

In our study recommended thresholds put more emphasis on sensitivity and positive predictive value, rating the disutility of overtreatment slightly smaller than the disutility of a missed case i.e. undertreatment.

Future

Non-clinical factors are important in the prediction of SGA. Further research should further explore the screening potential of the R4U. We expect a considerable improvement of the diagnostic test characteristics if the R4U is applied in an unselected group at the first booking visit. Its use in antenatal care could have some distinct advantages as it enables

early intervention and intensified follow-up. To improve the screening of SGA, the R4U could be used in combination with biochemistry and/or the use of a growth ultrasound. The R4U could as well be used for the prediction of other adverse outcomes.

SUMMARY

The R4U is a promising adjuvant diagnostic tool alongside the current risk selection. Non clinical risk factors should play a key role in the prediction of pregnant women at risk of SGA.

REFERENCES

1. Bonsel GJ, Birnie E, Denktas S, Poeran J, Steegers EAP. Dutch report: Lijnen in de Perinatale Sterfte, Signalementstudie Zwangerschap en Geboorte 2010 (http://www.nvk.nl/Nieuws/Dossiers/DossierPerinatalezorg.aspx). Rotterdam: Erasmus MC; 2010.
2. de Graaf JP, Ravelli AC, de Haan MA, Steegers EA, Bonsel GJ. Living in deprived urban districts increases perinatal health inequalities. J Matern Fetal Neonatal Med. 2012 Oct 8.
3. Poeran J, Denktas S, Birnie E, Bonsel GJ, Steegers EA. Urban perinatal health inequalities. J Matern Fetal Neonatal Med. 2011 Apr;24(4):643-6.
4. Ravelli AC, Tromp M, van Huis M, et al. Decreasing perinatal mortality in The Netherlands, 2000- 2006: a record linkage study. J Epidemiol Community Health. 2009 Sep;63(9):761-5.
5. Timmermans S, Bonsel GJ, Steegers-Theunissen RP, et al. Individual accumulation of heterogeneous risks explains perinatal inequalities within deprived neighbourhoods. Eur J Epidemiol. 2011 Feb;26(2):165-80.
6. CVZ. CV. List of Obstetric Indications (in Dutch: Verloskundig Vademecum) 2003. Diemen (http://www.knov.nl/docs/uploads/Verloskundig_Vademecum_2003.pdf). 2003.
7. Lindqvist PG, Molin J. Does antenatal identification of small-for-gestational age fetuses significantly improve their outcome? Ultrasound Obstet Gynecol. 2005 Mar;25(3):258-64.
8. van der Kooy J, Poeran J, de Graaf JP, et al. Planned home compared with planned hospital births in the Netherlands: intrapartum and early neonatal death in low-risk pregnancies. Obstet Gynecol. 2011 Nov;118(5):1037-46.
9. Raatikainen K, Heiskanen N, Heinonen S. Marriage still protects pregnancy. Bjog. 2005 Oct;112(10):1411-6.
10. Agyemang C, Vrijkotte TG, Droomers M, van der Wal MF, Bonsel GJ, Stronks K. The effect of neighbourhood income and deprivation on pregnancy outcomes in Amsterdam, The Netherlands. J Epidemiol Community Health. 2009 Sep;63(9):755-60.
11. Mortensen LH, Diderichsen F, Arntzen A, et al. Social inequality in fetal growth: a comparative study of Denmark, Finland, Norway and Sweden in the period 1981-2000. J Epidemiol Community Health. 2008 Apr;62(4):325-31.
12. Beard JR, Lincoln D, Donoghue D, et al. Socioeconomic and maternal determinants of small-for-gestational age births: patterns of increasing disparity. Acta Obstet Gynecol Scand. 2009;88(5):575-83.
13. Breeze AC, Lees CC. Prediction and perinatal outcomes of fetal growth restriction. Semin Fetal Neonatal Med. 2007 Oct;12(5):383-97.
14. de Graaf JP, Ravelli AC, Wildschut HI, et al. [Perinatal outcomes in the four largest cities and in deprived neighbourhoods in The Netherlands]. Ned Tijdschr Geneeskd. 2008 Dec 13;152(50):2734-40.

15. Fang J, Madhavan S, Alderman MH. Low birth weight: race and maternal nativity—impact of community income. Pediatrics. 1999 Jan;103(1):E5.

16. Gardosi J. Intrauterine growth restriction: new standards for assessing adverse outcome. Best Pract Res Clin Obstet Gynaecol. 2009 Dec;23(6):741-9.

17. Denktas S, Bonsel GJ, Van der Weg EJ, et al. An Urban Perinatal Health Programme of Strategies to Improve Perinatal Health. Matern Child Health J. 2011 Aug 26.

18. Goedhart G, van Eijsden M, van der Wal MF, Bonsel GJ. Ethnic differences in term birthweight: the role of constitutional and environmental factors. Paediatr Perinat Epidemiol. 2008 Jul;22(4):360-8.

19. Jaddoe VW, Verburg BO, de Ridder MA, et al. Maternal smoking and fetal growth characteristics in different periods of pregnancy: the generation R study. Am J Epidemiol. 2007 May 15;165(10):1207-15.

20. Timmermans S, Jaddoe VW, Hofman A, Steegers-Theunissen RP, Steegers EA. Periconception folic acid supplementation, fetal growth and the risks of low birth weight and preterm birth: the Generation R Study. Br J Nutr. 2009 Sep;102(5):777-85.

21. McCowan LM, Roberts CT, Dekker GA, et al. Risk factors for small-for-gestational-age infants by customised birthweight centiles: data from an international prospective cohort study. Bjog. 2010 Dec;117(13):1599-607.

22. Chauhan SP, Magann EF. Screening for fetal growth restriction. Clin Obstet Gynecol. 2006 Jun;49(2):284-94.

23. Romo A, Carceller R, Tobajas J. Intrauterine growth retardation (IUGR): epidemiology and etiology. Pediatr Endocrinol Rev. 2009 Feb;6 Suppl 3:332-6.

24. McCowan L, Horgan RP. Risk factors for small for gestational age infants. Best Pract Res Clin Obstet Gynaecol. 2009 Dec;23(6):779-93.

25. Nguyen D, Smith LM, Lagasse LL, et al. Intrauterine growth of infants exposed to prenatal methamphetamine: results from the infant development, environment, and lifestyle study. J Pediatr. 2010 Aug;157(2):337-9.

26. Meyer S, Raisig A, Gortner L, Ong MF, Bucheler M, Tutdibi E. In utero tobacco exposure: the effects of heavy and very heavy smoking on the rate of SGA infants in the Federal State of Saarland, Germany. Eur J Obstet Gynecol Reprod Biol. 2009 Sep;146(1):37-40.

27. Ananth CV, Kaminsky L, Getahun D, Kirby RS, Vintzileos AM. Recurrence of fetal growth restriction in singleton and twin gestations. J Matern Fetal Neonatal Med. 2009 Aug;22(8):654-61.

28. Verweij AvdL, F. Droomers, M. Rijksinstituut voor Volksgezondheid en Milieu (RIVM), Gezondheid in veertig krachtwijken. http://www.rivm.nl/bibliotheek/rapporten/270951001.pdf 2008.

29. Visser GH, Eilers PH, Elferink-Stinkens PM, Merkus HM, Wit JM. New Dutch reference curves for birthweight by gestational age. Early Hum Dev. 2009 Dec;85(12):737-44.

30. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. Bmj. 2009;338:b2393.

31. Seed PT, Chappell LC, Black MA, et al. Prediction of preeclampsia and delivery of small for gestational age babies based on a combination of clinical risk factors in high-risk women. Hypertens Pregnancy. 2011;30(1):58-73.

32. Grimes DA, Schulz KF. Uses and abuses of screening tests. Lancet. 2002 Mar 9;359(9309):881-4.

33. Troe EJ, Raat H, Jaddoe VW, et al. Explaining differences in birthweight between ethnic populations. The Generation R Study. Bjog. 2007 Dec;114(12):1557-65.

34. Drooger JC, Troe JW, Borsboom GJ, et al. Ethnic differences in prenatal growth and the association with maternal and fetal characteristics. Ultrasound Obstet Gynecol. 2005 Aug;26(2):115-22.

35. Alexander GR, Kogan MD, Himes JH, Mor JM, Goldenberg R. Racial differences in birthweight for gestational age and infant mortality in extremely-low-risk US populations. Paediatr Perinat Epidemiol. 1999 Apr;13(2):205-17.

36. Mor JM, Alexander GR, Kogan MD, Kieffer EC, Ichiho HM. Similarities and disparities in maternal risk and birth outcomes of white and Japanese-American mothers. Paediatr Perinat Epidemiol. 1995 Jan;9(1):59-73.

37. Schwartz N, Pessel C, Coletta J, Krieger AM, Timor-Tritsch IE. Early biometric lag in the prediction of small for gestational age neonates and preeclampsia. J Ultrasound Med. 2011 Jan;30(1):55-60.

38. Thiebaugeorges O, Fresson J, Audibert F, Guihard-Costa AM, Frydman R, Droulle P. Diagnosis of small-for-gestational-age fetuses between 24 and 32 weeks, based on standard sonographic measurements. Ultrasound Obstet Gynecol. 2000 Jul;16(1):49-55.

39. Bahado-Singh R, Oz U, Flores D, Hsu C, Mari G, Cole L. Maternal urine beta-core hCG fragment level and small for gestational age neonates. Obstet Gynecol. 2000 May;95(5):662-6.

40. Wen SW, Tan H, Yang Q, Walker M. Prediction of small for gestational age by logistic regression in twins. Aust N Z J Obstet Gynaecol. 2005 Oct;45(5):399-404.

41. Zhang H, Bracken MB. Tree-based risk factor analysis of preterm delivery and small-for-gestational-age birth. Am J Epidemiol. 1995 Jan 1;141(1):70-8.

42. NICE. Pregnancy and complex social factors. A model for service provision for pregnant women with complex social factors. London: National Institute for Health and Clinical Excellence. 2010;CG 110.

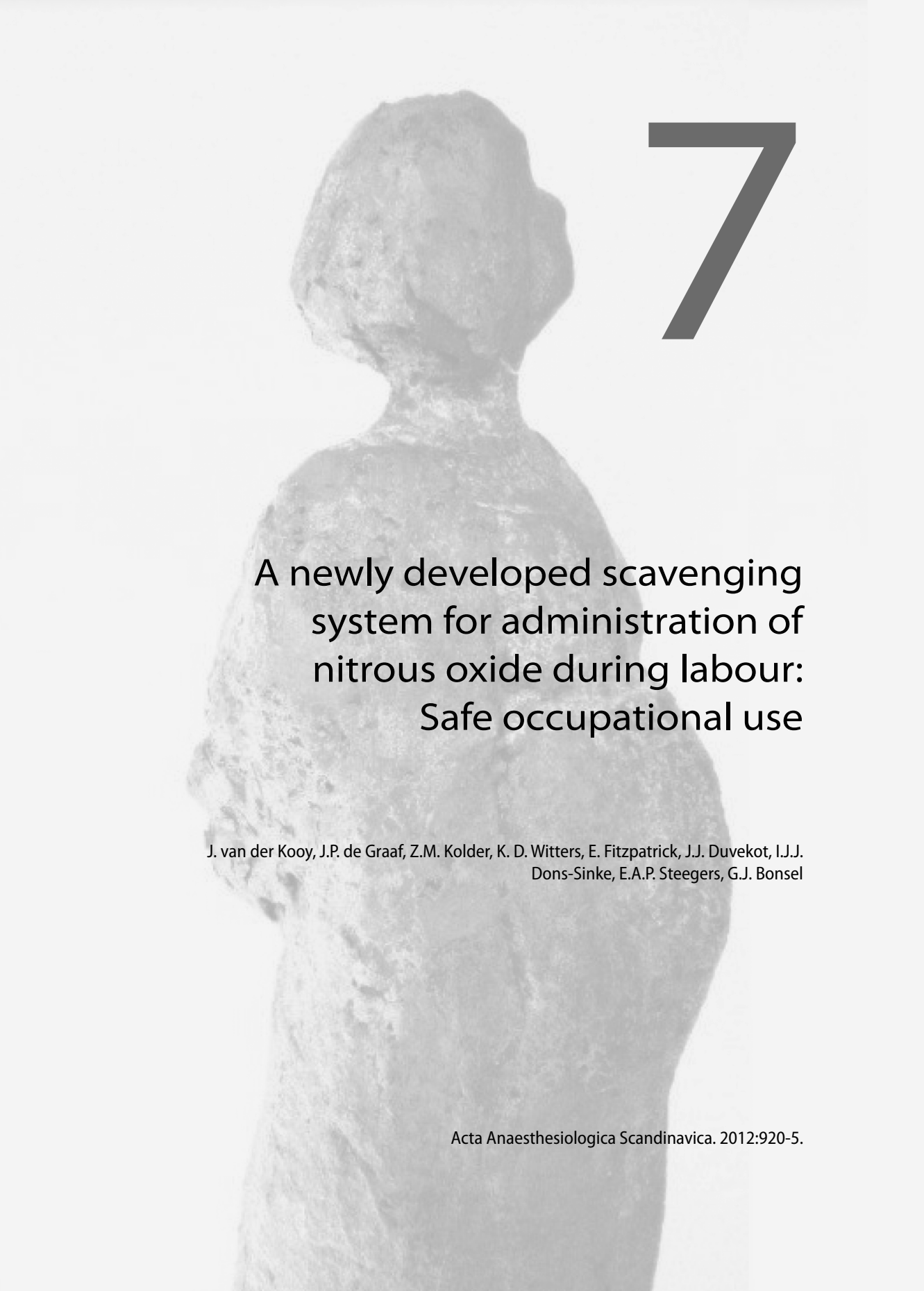
43. Rours GI, Duijts L, Moll HA, et al. Chlamydia trachomatis infection during pregnancy associated with preterm delivery: a population-based prospective cohort study. Eur J Epidemiol. 2011 Jun;26(6):493-502.

44. Bennini JR, Marussi EF, Barini R, Faro C, Peralta CF. Birth-weight prediction by two- and three-dimensional ultrasound imaging. Ultrasound Obstet Gynecol. 2010 Apr;35(4):426-33.

45. Dudley NJ. A systematic review of the ultrasound estimation of fetal weight. Ultrasound Obstet Gynecol. 2005 Jan;25(1):80-9.

46. Scioscia M, Scioscia F, Vimercati A, et al. Estimation of fetal weight by measurement of fetal thigh soft-tissue thickness in the late third trimester. Ultrasound Obstet Gynecol. 2008 Mar;31(3):314-20.

47. Alfirevic Z, Devane D, Gyte GM. Continuous cardiotocography (CTG) as a form of electronic fetal monitoring (EFM) for fetal assessment during labour. Cochrane Database Syst Rev. 2006(3):CD006066.



7

A newly developed scavenging system for administration of nitrous oxide during labour: Safe occupational use

J. van der Kooy, J.P. de Graaf, Z.M. Kolder, K. D. Witters, E. Fitzpatrick, J.J. Duvekot, I.J.J. Dons-Sinke, E.A.P. Steegers, G.J. Bonsel

Acta Anaesthesiologica Scandinavica. 2012:920-5.

ABSTRACT

Objective Nitrous oxide as an analgesic (N_2O) is routinely used in obstetrics during labour. Epidemiological studies have linked chronic occupational exposure to nitrous oxide to specific health problems, including reproductive risks. Occupational exposure limits allow the use of N_2O once appropriate preventive and safety measures have been taken. We assessed the effectiveness of a scavenger system (Anevac P-system®) applied in N_2O administration during labour in a midwifery led birthing centre in the Netherlands.

Methods After informed consent, non-pregnant midwives were trained to administer N_2O. N_2O was delivered as a 50:50 mixture with oxygen, and was self administered by the patient. The scavenging device, containing a double mask and a chin mask was connected to the local evacuation system vented outside the building. Data on the 8-hour Time Weighted Average (8-hr TWA) as well as the 15 minutes TWA (15-min TWA) was obtained.

Results Thirteen patients were included. Six patients were included in the first study period. In this period the 8-hour TWA was not exceeded, however in all patients the 15-min TWA occasionally exceeded the OELs. After four additional measures were taken, seven patients were included. After implementation of these measures the 8-hr TWA and 15-min TWA never exceeded the OELs. System leakage was not observed during both study periods.

Conclusion The Anevac P-scavenging system during N_2O analgesia in labour prevents exceeding occupational exposure limits in professional workers. The scavenging system appeared acceptable and effective, and can be considered in hospital settings that use N_2O as analgesic during labour.

INTRODUCTION

Nitrous oxide as analgesic (N_2O) is routinely used in dentistry, in the emergency room, and during labour. It combines several advantages compared to intravenous alternatives like propofol or, in obstetrics epidural analgesia, which makes it a preferred option in specific contexts.[1-2] Internationally a wide variety in N_2O application as well as a wide variety in imposed safety regulations is observed. This may reflect uncertainty about how to apply N_2O in a truly safe manner.

Epidemiological studies have linked long term occupational exposure to nitrous oxide with reproductive risks such as spontaneous abortion, congenital anomalies and reduced rates of fertility.[3-12] Also adverse effects on the haematological and nervous system have been described.[4,13-16]

The recognition of the potential hazard to professional workers who are routinely exposed to nitrous oxide, elicited the introduction of occupational exposure limits (OELs) in different countries.[17] Governmental legislation enforces adherence to these OELs.

An OEL is expressed as a health-based occupational exposure limit, an eight hour time weighted average (8-hr TWA). In some countries a short term exposure limit, 15-minutes TWA (15-min TWA), has been adopted. In the Netherlands the 8-hr TWA for nitrous oxide is set on 152 mg/ m^3 whereas the 15-min TWA is set on 304 mg/m^3. Studies have already shown that midwives are regularly exposed to higher levels of nitrous oxide than permitted.[18-20] Futhermore studies showed that the need for anaesthetic waste gas scavenging is of great importance in order not to exceed the OELs.[20 21-23] Strict University Hospital regulations in the Netherlands demand exposure levels to medical professionals which are less than or equal to only 25% of the recommended OELs (38 mg/m^3 resp. 76 mg/m^3). The use of nitrous oxide has declined within obstetric care in the Netherlands since no scavenger system is available in standard settings.

In this study we assessed the practical appliance, patient convenience and effectiveness of a scavenger system (Anevac P-system®, Medicvent Heinen & Löwestein Benelux, Barneveld, The Netherlands) given to women who received nitrous oxide during labour in a midwifery led birth centre in the Netherlands. This study was undertaken to apply for approval of the Dutch National Institute for Occupational Safety and Health.

METHODS

The study protocol was approved by Medical Ethics Committee. This prospective observational intervention study was conducted during October 2009 and February 2010.

All midwives (n=15) working in a midwifery led birth centre were invited to participate and were asked to provide their written consent, unless they were pregnant or possibly pregnant or suffering from a known vitamin B deficiency. After informed consent was given they were trained in the administration of nitrous oxide and the use of the new scavenger system.

During the study period patients were informed about the option for nitrous oxide analgesic (N_2O) during labour and were asked to provide verbal consent. They were not permitted to use N_2O when acquaintances attending the birth were pregnant or possibly pregnant. In that case they were offered another form of analgesia and were referred to the gynaecologist. If anyone attending the delivery was suffering from a vitamin B12 deficiency they were excluded from the study because of possible adverse effects on the haematological system.

Nitrous oxide analgesic was administrated as a 50:50 mixture with oxygen, know as Relivopan®. It was delivered from a portable N_2O gascilinder using a pin index system [Linde Gas Benelux, Schiedam, the Netherlands]. Nitrous oxide was self administered by the patient through a double mask containing a Carnét Demand valve [Medicvent, Heinen & Löwenstein, Groningen, the Netherlands].

The scavenging device, contained a double mask and a chin mask (Anevac P-system®) [Medicvent, Heinen & Löwenstein, Groningen, Nederland]; it was connected to the local evacuation system vented outside the building. The evacuation rate of the scavenging device was tested with an in-line flowmeter and was found to produce 34 m^3/hr (17 m^3/hr through the double mask, 17 m^3/hr through the chin mask). (figure 7.1)

Room air exchange rates were 6 air changes per hour for each individual delivery room. The patient was instructed to use the double mask and chinmask during the first stage of labour. Entering the second stage of labour both the double mask and chin mask were removed, after discontinuation of N_2O administration. During the first stage of labour the use of the scavenging system was thoroughly observed to identify risk factors for possible leakage. Apart from the initial instruction little correction during administration was given, to obtain results close to practice.

Data on the 8-hr TWA as well as information on 15-min TWA was obtained from the start of N_2O administration in the first stage till the third stage of labour. This was done to investigate N_2O levels while administrating and after discontinuation of administration. For the 8-hr TWA exposure a nitrous oxide diffusion sampler was clipped to the midwives lapel to keep it as close to the breathing area as possible. Diffusion samplers are generally designed to sample over a period of time for the determination of average concentrations. In contrast to active sampling the transport of the contaminant molecules is achieved by diffusion

Figure 7.1 Scavenging device (photo with written permission of the patient).
(1) Demand valve
(2) Double mask
(3) Chin mask

processes and not by using a pump. Contaminants from ambient air are adsorbed by the sorption agent (www.draeger.com). After sampling the tubes were analysed with Gas Chromatography analysis and Mass-Spectrometry detection. They were asked to record the length of the shift and how long they spent on the labour suite. The measurement of the 15-min TWA was obtained by the 1312 Photoacoustic Multi-gas Monitor (detection limit for nitrous oxide is 0.06 mg/m^3) which absorbs an air sample once every 60 seconds and directly analyses the concentration nitrous oxide. This sensitive technique allows for direct measurement of high risk acts and system leakage. This technique was in permanent use during all observations. (For a more detailed description see: http://www.lumasenseinc.com/EN/products/gas-monitoring-instruments/gas-monitoring/technical-information-of-gas/photoacoustic-detection-pas.html)

Patients as well as health care workers/professionals were asked to fill out a questionnaire about the practical appliance and convenience of the administration and scavenger system.

Interpretation of the exposure assessment was based upon 25% of the set OELs. Subsequently we used the cut off value of 38 mg/m^3 (20 ppm) resp. 76 mg/m^3 (39 ppm).

Having analysed the results of the first five patients additional improvements to ensure maximum system effectiveness were performed in the second study period in February 2010. These included; (1) discontinuation of N_2O if the patient continued to be restless after 15 minutes, (2) permanent adequate position of chin mask during first and for at least 20 minutes in the second stage of labour, (3) extra 100% oxygen for 5 minutes administered after discontinuation N_2O administration, when entering the second stage of labour and (4) throughout second stage increased evacuation rate (34 L/min) of the chin mask.

RESULTS

In October 2009 six patients were included. One patient was excluded because nitrous oxide was only given for 15 minutes. About 23 patient hours were continuously and intensely monitored; in 19 of 23 hours nitrous oxide was actually administered. Analyses are shown in table 7.1. In this study period the 8-hr TWA was not exceeded, however in all patients short term peaks were observed. The 15-min TWA was exceeded when the chin mask was either not accurately positioned or was removed by the restless patient. Exceeding of the 15-min TWA were also noted during second stage after the discontinuation of N_2O when the chin mask was removed in agreement with protocol

Table 7.1 Results of the first study period in October 2009

Patient	Total time N_2O administration (minutes)	Average occupational exposure (mg/m^3)	8-hour TWA (mg/m^3)	15- minutes TWA (mg/m^3) only given when exceeding the OELs			Comments
1	180	35	13	170	231		Chin mask was not accepted. Referred to gynaecologist for other analgesia.
2	275	20.2	12	115	90	161	Chin mask removed when entering second stage of labour.
3	250	13.5	7	170			Chin mask removed when entering second stage of labour.
4	195	38.2	16	214			Patient restless, extremely vocal. Referred to gynaecologist for other analgesia.
5	207	31.5	14	101			Chin mask removed when entering second stage of labour.

in the first study period, since the protocol prescribed the use of the scavenging system only while administrating N_2O.

After introducing the above described improvements seven more patients were included in February 2010. Two of these were excluded from the analysis because they remained restless after 15 minutes of nitrous oxide administration. In this period a total of 21 hours (over 5 patients) were continuously observed. Within these 21 hours nitrous oxide was given for 16 hours. In the second study period the 8-hr TWA as well as the 15- min TWA were not exceeded. In this period one control patient was included, in whom no scavenging was used. In this patient both the 8-hr TWA and the 15-min TWA substantially exceeded the OEL's. Analyses are shown in table 7.2.

No system leakage was found during both study periods.

Equipment was found to be user friendly by both patients and caregivers. After instruction to the patient the 'on demand' administration of nitrous oxide was found to be convenient. Only one patient removed the chin mask. The remaining patients showed no discomfort while using the mask.

Table 7.2 Results of the second study period in February 2010

Patient	Total time N_2O administration (minutes)	Average occupational exposure (mg/m^3)	8-hour TWA (mg/m^3)	15 minutes TWA (mg/m^3) only given when exceeding the OELs	Comments
7	186	12	5	No peaks observed	Oxygen for 5 min and chin mask continually worn.
8	250	26.9	14	No peaks observed	Oxygen for 5 min and chin mask continually worn.
9	195	23.8	10	No peaks observed	Oxygen for 5 min and chin mask continually worn.
10	135	21.7	6	No peaks observed	Oxygen for 5 min and chin mask continually worn.
					Referred to gynaecologist due to failure to progress.
11	188	15.9	6	No peaks observed	Oxygen for 5 min and chin mask continually worn.
12	203	1,582	663	Concentrations exceeding 76 mg/m^3 continually	Control patient, no scavenging used.

DISCUSSION

This study stresses the importance of using a scavenging system containing a doublemask and chinmask and applying it with the four additional improvements to improve system effectiveness. It shows that the 8-hr TWA and the 15-min TWA are met by the use of the Anevac P-scavenging system while administrating nitrous oxide during labour. In addition, the introduction of four additional improvements increased system effectiveness. Limitations of the study include that the Anevac P-scavenging system was tested in the continueous presence of a Health, Safety and Environment specialist. This may lead to an higher compliance to the study protocol. Protocol compliance should therefore be watched closely during and after implementation. A possible second limitation is the relative small number of patients included. Despite this, observations were performed thoroughly including a sensitive technique which allowed for direct measurement of high risk acts for the spilling of nitrous oxide. The results of our qualitative evaluation of the Multigas monitor, as well as the systemic evaluation and documentation suggest that the administration of nitrous oxide should be discontinued when the patient remains restless after 15 minutes or when no adequate position of chin mask is reached. To realize the expected effect of nitrous oxide a small 'run in' period of 15 minutes is needed, due to the instruction to the patient and the correct procedure to administer nitrous oxide. When this state is not reached within 15 min, it is highly unlikely that it will with a longer 'run in' period, and is therefore likely to exceed the OELs.

After entering second stage of labour, saturated oxygen (100%) should be inhaled for 5 minutes and the chin mask should be worn continuous, with an increased evacuation rate (34 L/min). Postoperative oxygen is commonly given to prevent hypoxemia in patients and to wash out anaesthetic gases.[24]

Studies showed that the need for anaesthetic waste gas scavenging during labour is of great importance in order not to exceed OELs.[20-23]

Heath et all.[20] found average concentrations of 52 mg/m^3 with the use of a scavenging system compared to 297 mg/m^3 where none was used. Munley et al[22] found the same results seeing to lower exposure levels when a scavenging system was used.

In our study the 8-hr TWA concentrations ranged between 12,0 mg/m^3 and 38,2 mg/m^3. This it relatively low compared to others. Westberg et al [21] found concentrations ranging between 2,5 mg/m^3 and 260 mg/m^3. After comparing a simple facemask with a double mask they considered that their results favor the use of the double mask. The same result is seen by Chessor et al.[23]

Comparing the use of a simple face mask with double mask the observed exposure levels varied between 40mg/m^3 and 216 mg/m^3 (average 125 mg/m^3) and 10mg/m^3 and 306 mg/m^3 (average 72 mg/m^3) respectively. In addition they highlighted one of the main difficulties with scavenging systems when used in labor and delivery, being under patient control, the mask is most frequently held at a distance too far from the face to allow scavenging of exhaled breath. Both these result confirm our findings on the use of the double mask and the relevance of the use of the chinmask. Newton et al[18] compared exposure levels in 2 buildings; an older building where levels varied between 32 mg/m^3 and 2071 mg/m^3 and in a more modern facility comparable to ours, where levels varied between 14 and 172 mg/m^3. This study might underline the importance of a good ventilation system.

Despite the use of scavenging systems, OELs are still exceeded in these studies.[20-23] This stresses the importance of using a scavenging system containing a doublemask and chinmask and applying it with the four additional improvements to improve system effectiveness. Midwives must be trained regarding these improvements, understanding the reasons for implementation. To ensure the use of nitrous oxide as a useful analgesic during labour can be continued, scavenging equipment is required. This scavenging system turns out to be practical and effective and should therefore be considered in clinics that use nitrous oxide during labour.

After presenting the data approval was granted by the Dutch National Institute for Occupational Safety and Health.

REFERENCES

1. Harrison RF, Shore M, Woods T, Mathews G, Gardiner J, Unwin A. A comparative study of transcutaneous electrical nerve stimulation (TENS), entonox, pethidine + promazine and lumbar epidural for pain relief in labor. Acta Obstet Gynecol Scand. 1987;66(1):9-14.
2. Rosen MA. Nitrous oxide for relief of labor pain: a systematic review. Am J Obstet Gynecol. 2002 May;186(5 Suppl Nature):S110-26.
3. Baird PA. Occupational exposure to nitrous oxide—not a laughing matter. N Engl J Med. 1992 Oct 1;327(14):1026-7.
4. Cohen EN, Gift HC, Brown BW, et al. Occupational disease in dentistry and chronic exposure to trace anesthetic gases. J Am Dent Assoc. 1980 Jul;101(1):21-31.
5. Doll R, Peto R. Mortality among doctors in different occupations. Br Med J. 1977 Jun 4; 1(6074):1433-6.
6. Ericson A, Kallen B. Survey of infants born in 1973 or 1975 to Swedish women working in operating rooms during their pregnancies. Anesth Analg. 1979 Jul-Aug;58(4):302-5.
7. Kugel G, Norris LH, Zive MA. Nitrous oxide and occupational exposure: it's time to stop laughing. Anesth Prog. 1989 Nov-Dec;36(6):252-7.
8. McGregor DG, Senjem DH, Mazze RI. Trace nitrous oxide levels in the postanesthesia care unit. Anesth Analg. 1999 Aug;89(2):472-5.
9. Rowland AS, Baird DD, Weinberg CR, Shore DL, Shy CM, Wilcox AJ. Reduced fertility among women

employed as dental assistants exposed to high levels of nitrous oxide. N Engl J Med. 1992 Oct 1; 327(14):993-7.

10. Schumann D. Nitrous oxide anaesthesia: risks to health personnel. Int Nurs Rev.1990 Jan-Feb;37(1):214-7.

11. Spence AA. Environmental pollution by inhalation anaesthetics. Br J Anaesth. 1987 Jan;59(1):96-103.

12. Vessey MP. Epidemiological studies of the occupational hazards of anaesthesia--a review. Anaesthesia. 1978 May;33(5):430-8.

13. Middendorf PJ, Jacobs DE, Smith KA, Mastro DM. Occupational exposure to nitrous oxide in dental operatories. Anesth Prog. 1986 Mar-Apr; 33(2):91-7.

14. Sharer NM, Nunn JF, Royston JP, Chanarin I. Effects of chronic exposure to nitrous oxide on methionine synthase activity. Br J Anaesth. 1983 Aug;55(8):693-701.

15. Landon MJ, Creagh-Barry P, McArthur S, Charlett A. Influence of vitamin B12 status on the inactivation of methionine synthase by nitrous oxide. Br J Anaesth. 1992 Jul;69(1):81-6.

16. Axelsson G, Ahlborg G, Jr., Bodin L. Shift work, nitrous oxide exposure, and spontaneous abortion among Swedish midwives. Occup Environ Med. 1996 Jun;53(6):374-8.

17. Sanders RD, Weimann J, Maze M. Biologic effects of nitrous oxide: a mechanistic and toxicologic review. Anesthesiology. 2008 Oct;109(4):707-22.

18. Newton C, Fitz-Henry J, Bogod D. The occupational exposure of midwives to nitrous oxide - a comparison between two labour suites. Int J Obstet Anesth. 1999 Jan;8(1):7-10.

19. Mills GH, Singh D, Longan M, O'Sullivan J, Caunt JA. Nitrous oxide exposure on the labour ward. Int J Obstet Anesth. 1996 Jul;5(3):160-4.

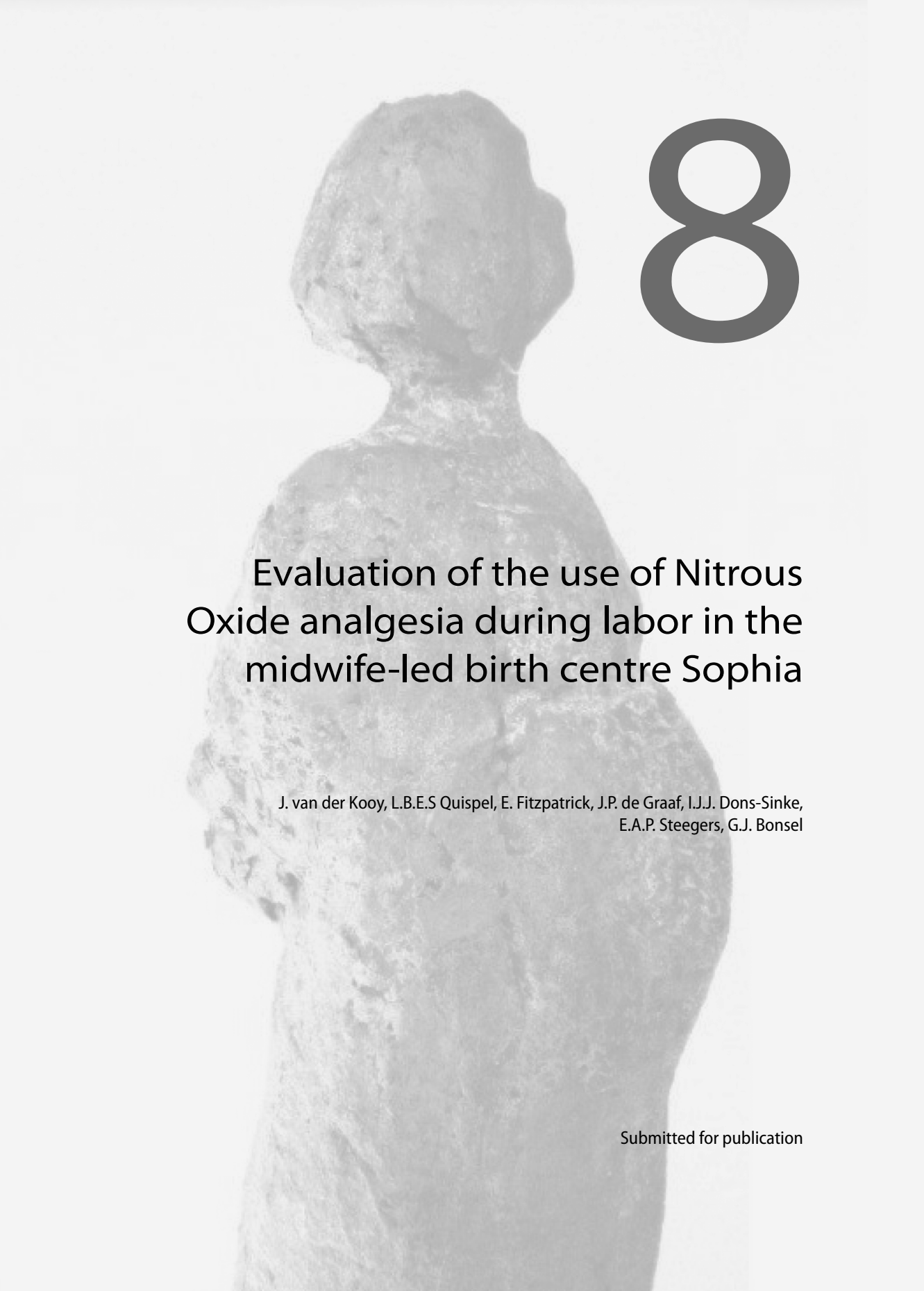
20. Heath BJ, Done M, Balog O, Ziccone S, Rosewarne F. The effect of scavenging on nitrous oxide pollution in the delivery suite. Aust N Z J Obstet Gynaecol. 1994 Aug;34(4):484-6.

21. Westberg H, Egelrud L, Ohlson CG, Hygerth M, Lundholm C. Exposure to nitrous oxide in delivery suites at six Swedish hospitals. Int Arch Occup Environ Health. 2008 Jul;81(7):829-36.

22. Munley AJ, Railton R, Gray WM, Carter KB. Exposure of midwives to nitrous oxide in four hospitals. Br Med J (Clin Res Ed). 1986 Oct 25; 293(6554):1063-4.

23. Chessor E, Verhoeven M, Hon CY, Teschke K. Evaluation of a modified scavenging system to reduce occupational exposure to nitrous oxide in labor and delivery rooms. J Occup Environ Hyg. 2005 Jun;2(6):314-22.

24. Powell JF, Menon DK, Jones JG. The effects of hypoxaemia and recommendations for postoperative oxygen therapy. Anaesthesia. 1996 Aug;51(8):769-72.



8

Evaluation of the use of Nitrous Oxide analgesia during labor in the midwife-led birth centre Sophia

J. van der Kooy, L.B.E.S Quispel, E. Fitzpatrick, J.P. de Graaf, I.J.J. Dons-Sinke, E.A.P. Steegers, G.J. Bonsel

Submitted for publication

ABSTRACT

Background There is a rising need for pain relief during labor. Nitrous Oxide (N_2O) administration is considered a safe and effective option. The aim of this prospective observational study is to evaluate whether N_2O analgesia is a valid option within a birthing centre adjacent to the hospital.

Methods We selected the records of all singleton low risk pregnant women who started their delivery in the birthing centre Sophia between January 2010 and December 2011 (n=1437). N_2O analgesia was introduced in January 2011 and was administrated as a 50:50 mixture with oxygen using a standardized protocol. N_2O was self administered by the patient through a double mask. Outcome measurements of side effects and analgesic effectiveness of N_2O were provided independently by the midwives and women.

Results The introduction of N_2O in the birth centre Sophia led to an overall 8% increase in the use of analgesia in labor, which translates to 160 out of 727 women (22%, 95%CI 19-25%) before and 211 out of 710 (30%, 95%CI 26-33%) after (p=0.002). The proportion of women receiving other analgesia (pethidine and/or epidural) was reduced by 5% (p=0.02). Interestingly, this trend was mainly observed in multiparous women aged 25-34 years. Scores on pain, exhaustion and anxiety decreased during labor in women receiving N_2O analgesia. However, after starting N_2O analgesia, 16 (15%) out of 106 women needed additional analgesia.

Conclusions From this single centre study we conclude that N_2O analgesia is worthwhile to be considered as an analgesic option within birth centers adjacent to hospitals.

INTRODUCTION

There is a rising need for pain relief during labor.[1] Several analgesia options are available, such as epidural analgesia (EDA), the administration of remifentanil, pethidine and Nitrous Oxide analgesia (N_2O). EDA is found to be most effective which makes EDA the clinically preferred option.[2,3] The question has emerged if other options, in particular N_2O analgesia, are still useful, as N_2O analgesia is less invasive.[4,5]

Labor pain is a subjective response that is influenced by a woman's beliefs, expectations and values, and her environment.[6] Consequently, it is worthwhile to investigate less costly or less invasive alternatives like N_2O analgesia. While it may on average be less effective than EDA, it could be sufficient for certain subgroups of women under particular conditions.[3,4] Advantages of N_2O analgesia compared to other analgesic options are its immediate and intermittent availability, being non invasive, and low costs.[5,7]

N_2O analgesia is mainly used in Scandinavian and Anglo-Saxon countries during labor at this stage and less so in other countries.[8] N_2O analgesia use in the Netherlands has strongly declined, since thorough measures are needed to reach occupational safety.[9]

The evidence of the analgesic effect of N_2O analgesia is limited, as studies were small, data were collected retrospectively, and as measurements focus on levels of pain rather than the need for additional pain relief. Safety aspects, in particular O2 desaturation was not systemically addressed.[10-12] In addition, when evaluating effectiveness one focused purely upon its analgesic effects. However, surveys suggest that around 80% of mothers find N_2O analgesia helpful and would use it in a future delivery suggesting that although N_2O analgesia provides much less complete pain relief than EDA, for many women, it is enough.[3,8,12]

In the Netherlands N_2O analgesia is considered a sufficient option for midwife-led birth centers adjacent to a hospital. In the Dutch system, independently operating community midwives provide the care for women with low risk pregnancies (primary healthcare). These women can either deliver at home, in hospital or in a midwife-led birth centre under sole supervision of the midwife. If N_2O analgesia is successful, the need for referral for pain relief to a gynecologist led obstetric unit is avoided. Women with high-risk pregnancies, as defined in the Dutch guidelines[13] (e.g. history of postpartum hemorrhage, presence of a hypertensive disorder, presence of small for gestational age) do not have the choice of delivering in the birthing centre but receive ante- and intrapartum care in the hospital under the supervision of the gynecologist (secondary healthcare). During our research N_2O was not available in the hospital yet.

The aim of this prospective observational study is to evaluate whether N_2O analgesia is a valid option within a midwife led birth centre adjacent to a hospital. We examined the frequency of use of N_2O analgesia during labor and whether its use precluded the use of other analgesia, in our case pethidine and/or epidural. Furthermore, we looked at N_2O efficacy as a form of pain relief and overall patient satisfaction. In addition, the occurrence of complications was assessed.

METHODS

Data collection

Our department of Obstetrics and Gynecology is responsible for independent permanent monitoring of the efficiency and safety of N_2O analgesia in the midwife-led birth centre Sophia. This centre Sophia is build adjacent to our university hospital wards and offers care to low risk women during labor under the supervision of a community midwife. This includes women who start birth between 37 and 42 weeks of gestation with no medical or obstetric history as defined in the Dutch guidelines.[13] During labor obstetric, neonatal and anesthetic care remained readily available in the adjacent hospital should the need for referral arise.

The midwife-led birth centre was licensed to administrate N_2O analgesia by a trained midwife, following a predefined protocol (see below).[14] Standardized questionnaires were routinely used during and after delivery. Additional data was drawn from the patient records and were accessible for research purposes under standardized limitations.

We selected the records of all singleton pregnant women, under supervision of a community midwife, who entered the birth centre Sophia at the onset of labor between January, 01 and December, 31 2011. We excluded records were the data was incomplete (n=11). For comparison we selected records of all singleton pregnant women entering the birth centre Sophia at the onset of labor before the introduction of Nitrous Oxide (January, 1 until December, 31 2010). No personal data were abstracted (strict anonymous data). Dutch law exempts institutional review of the Medical Ethics Committee since anonymized medical records were used retrospectively.

N_2O administration

N_2O analgesia was routinely offered to women requesting analgesia. A dual decision was taken by the midwife and the pregnant woman as to the choice of analgesia.

None of the women had received any analgesia before commencement of N_2O analgesia. In addition, N_2O was never given in combination with any other type analgesia. In order to apply N_2O analgesia the midwife made sure the patient did not have any exclusion criteria for N_2O analgesia, including: increased intracranial pressure, pneumothorax, pneumopericardium, chronic obstructive lung disease, emphysema, history of heart surgery, vitamin B12 deficiency.

If N_2O was insufficient in terms of analgesia, a switch was made to an alternative in the obstetric unit (pethidine and/or EDA) and the administration of N_2O analgesia was thus discontinued.

All midwives had been trained to administer N_2O analgesia using a standardized protocol.[14] This protocol included instructions on the timing, administration and monitoring of N_2O during labor. In addition, the protocol included instructions on the use of a newly developed scavenging system [Anevac P-system], to avoid chronic exposure of the caregivers to N_2O.

N_2O was administrated as a 50:50 mixture with oxygen, known as Relivopan®. It was delivered via a portable N_2O gas cylinder using a pin index system [Linde Gas Benelux, Schiedam, the Netherlands]. Nitrous Oxide was self administered by the patient through a double mask containing a Carnét Demand valve [Medicvent, Heinen & Löwenstein, Groningen, the Netherlands]. The scavenging system, consisted of a double mask and a chin mask (Anevac P-system®) [Medicvent, Heinen & Löwenstein, Groningen, Nederland]; it was connected to the local external evacuation system. The evacuation rate of the scavenging device was tested with an in-line flow meter and was found to produce 34 m^3/hr (17 m^3/hr through the double mask, 17 m^3/hr through the chin mask).[14]

Outcome data

Degree of pain was only measured during the administration of N_2O using a visual analogue scale (VAS score), consisting of a 100 mm horizontal line with a verbal description at either end; a score of 0 represents "no pain" and 100 represented "worst pain imaginable". VAS scores were obtained at baseline, 15-30 minutes and 60 minutes after start N_2O analgesia. The first interval (15-30 minutes) was chosen to measure the quick effect of N_2O after the women was appropriately instructed. This instruction often took about 2-3 contractions. The second interval (60 minutes) was chosen to assess the effect of N_2O administration over a longer period.

Each woman was additionally asked to fill in a self report questionnaire 24 hours after her delivery on her overall experience of N_2O. The woman was invited to rate her pain experience, exhaustion, tiredness and her anxiety, using a 5 point-Likert scale. She was asked to rate in retrospect her experience at; baseline, 15-30 minutes, and 60 minutes after N_2O administration.

During delivery, the midwife recorded her own assessment of the woman's pain experience, the exhaustion, the tiredness and the anxiety using a 5 point-Likert scale. This was prospectively recorded at baseline (before the start of N_2O), 15-30 minutes, and 60 minutes after N_2O administration. The midwife additionally recorded the presence of any side effects, such as nausea and vomiting. Outcomes were only obtained on N_2O administration and stopped when the woman were referred to secondary care.

Other medical data

In addition cervical dilatation, blood pressure, heart rate, and Ramsay score of sedation were recorded at baseline by the midwife. Ramsay score used the following categories; 1. awake and oriented, 2. sleepy and calm, 3. responsive to commands only, 4. responsive to physical stimulus only, 5. unresponsive.[15]

Periodic monitoring with a pulse oximeter and a blood pressure monitor was instituted using a Dinamap Carescape v100® [Medidis BV, Lelystad, the Netherlands]. Blood pressure, heart rate, saturation, Ramsay score of sedation and VAS were recorded throughout labor at 15, 30, and 60 minutes of N_2O administration. Duration of delivery was defined as the time between the start of contractions and the time of birth. This was recorded after the delivery.

Maternal characteristics, maternal outcomes and neonatal outcomes were derived from the medical records. The characteristics were derived in order to assess their influence on the choice of analgesic options. They included parity (primiparous vs. multiparous), age, ethnicity (Dutch/Non Dutch), neighborhood (deprived/privileged, based on 4-digit zip-codes and a public list of deprived neighborhoods based on zip codes issued by the Dutch government)[16] and the onset of second stage of labor defined as full cervical dilatation.

Statistical analysis

Data were entered into SPSS version 17.0.

Differences in maternal characteristics were analyzed for the period before and after the introduction of N_2O. The primary analysis was a pre-post comparison to assess for which

group of women the availability of N_2O analgesia averted the use of other analgesics (substitution). Furthermore, analysis was made of the group of women for whom the availability of N_2O led to an overall increase in the use of analgesia(additional use), those being pethidine and/or EDA.

Maternal characteristics were compared using standard Student's t-test or chi-squared analysis where appropriate. To assess the characteristics of women where N_2O analgesia averted the use of other analgesic options, the characteristics of women who received pethidine and/or EDA were compared between the two periods 2010 and 2011.To assess which women, after the introduction of N_2O analgesia, showed an overall increase in analgesic use we compared women who received no analgesia and delivered in 2010 with the women receiving no analgesia in 2011.

The absolute effectiveness of N_2O was obtained by examining the women who needed additional analgesia after starting N_2O analgesia. In addition, the absolute effectiveness of N_2O was obtained by examining the women who needed additional analgesia after starting N_2O analgesia. In addition, the absolute effectiveness of N_2O was obtained by assessing the differences between baseline and 15-30 minute outcome measurements. This time interval was chosen as best proxy for effect, since selection effects (drop out of worse cases) is thought to be less present compared to the 60 minute interval.

Outcome measurements included VAS scores and the percentage of women in the very/extreme category on the pain, exhaustion and anxiety Likert scales. We decided a difference of 10mm in VAS to be clinically relevant based on previous studies.[17-19] This difference is only calculated for those patients who provided outcome measurements for both time intervals. Statistical difference was measured using paired t-test. No statistical analysis was made for the Likert scales, since clinically relevant differences have yet to be identified.

Comparison on these outcomes was made between the group in which N_2O analgesia sufficed during labor (Nitrous Oxide only) with the group in which N_2O analgesia was not effective enough, thus leading to the need for additional analgesia (Nitrous Oxide and Pethidine and/or EDA). The presence of any side effects was given for the two periods.

RESULTS

The study population consisted of 1437 women. Table 8.1 describes the characteristics of women in the period before and after the introduction of N_2O analgesia. 727 women were included in the period before N_2O analgesia was introduced (2010) and 710 women in the

period after its introduction (2011). In the period before N_2O introduction, 392 (54%) were primiparous, 407 (56%) from a non Dutch origin, and 353 (49%) were from an unprivileged neighborhood. In the period after N_2O introduction, 396 (52%) were primiparous, 497 (56%) from a non Dutch origin, and 367 (52%) were from an unprivileged neighborhood. No statistical differences were observed between the two periods.

Table 8.1 Characteristics of women starting their labor in a midwife-led birth centre before and after the introduction of Nitrous Oxide

	Before the introduction of Nitrous Oxide analgesia (January - December 2010)		After introduction of Nitrous Oxide analgesia (January - December 2011)		
Variable	n	727 %	n	710 %	p-value
Parity					0.492
Primiparous	392	54%	369	52%	
Multiparous	335	46%	341	48%	
Maternal Age					0.125
<25 years	135	18%	149	21%	
25-34 years	438	60%	429	60%	
>35 years	148	20%	117	16%	
Missing	6	1%	15	2%	
Ethnic background					0.732
Dutch	256	35%	236	33%	
Non Dutch	407	56%	397	56%	
Missing	64	9%	77	11%	
Neighbourhood					0.246
Privileged neighbourhood	374	51%	343	48%	
Underprivileged neighbourhood	353	49%	367	52%	
Time entering second stage of labour*					0.218
0-8 hr	230	32%	246	35%	
8-18 hr	319	44%	254	36%	
18-24 hr	140	19%	144	20%	
Missing	38	5%	66	9%	

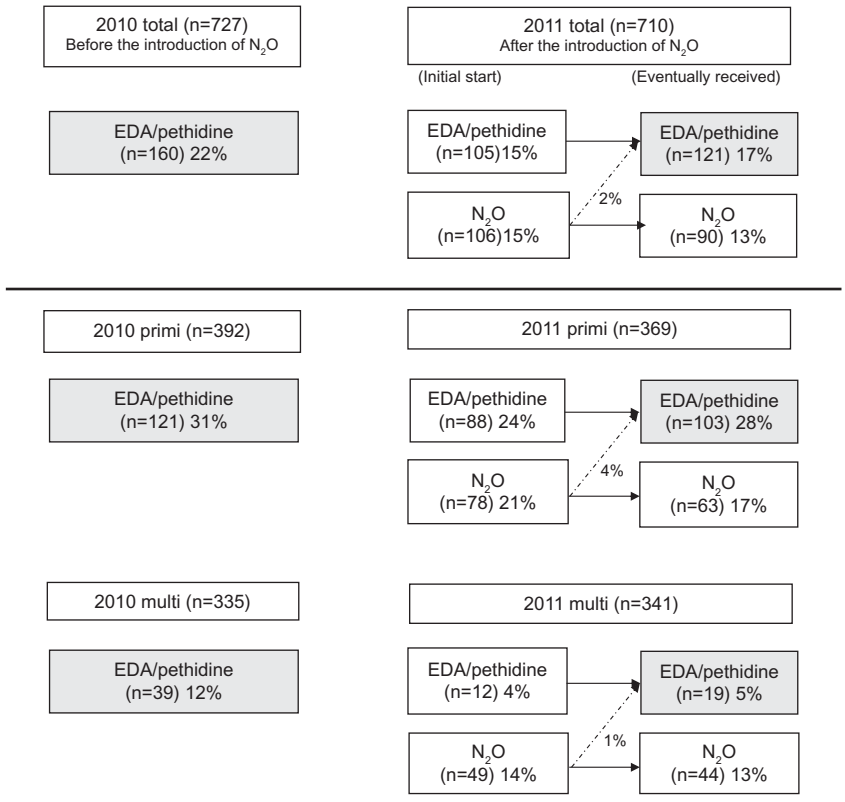


Figure 8.1 Analgesic utilization pattern for the periods 2010 (period before N_2O introduction) and 2011 (period after N_2O introduction).

Figure 8.1 compares the analgesic utilization pattern for the period 2010 and for the period 2011. This pattern is given for the total group and stratified for parity. In 2010 160 (22%, 95%CI 19-25%) of all women who started labor in the midwife-led birth centre Sophia ended up being transferred to the hospital to receive additional analgesia (pethidine and/or EDA). In 2011 211 (30%, 95%CI 26-33%) received analgesia (either pethidine, EDA or N_2O), a significant increase of 8%(p=0,002). The proportion of women receiving pethidine and/or EDA in 2011 was 17% (95%CI 14-19%), a significant reduction of 5% compared to the proportion in 2010 (p=0.023). After stratification for parity a similar increase in use and substitution was seen for both primi- and multiparous women. However, within the group of primiparous women a higher increase in the overall use of analgesia was observed (14% vs. 6%) compared to multiparous women, and less substitution took place (3% vs. 7%).

Table 8.2 describes the baseline characteristics of all women who gave birth within the midwife-led birth centre Sophia between January 2010 and December 2011, subdivided

Table 8.2 Characteristics of women for the periods before and after the introduction of Nitrous Oxide, subdivided for their analgesic use

	Before the introduction of Nitrous Oxide analgesia (January - December 2010)				After the introduction of Nitrous Oxide analgesia (January - December 2011)			
	A No analgesia		B First choice EDA/ pethidine		C No analgesia		D First choice Nitrous Oxide	
							D1 Nitrous Oxide only	
Variable	n=n	567 (78%)	n=n	160 (22%)	n=n	499 (70%)	n=n	90 (13%)
Column		%		%		%		%
Parity								
Primiparous	271	47.8%	121	75.6%	210	42.1%	55	61.1%
Multiparous	296	52.2%	39	24.4%	289	57.9%	35	38.9%
Maternal Age								
<25 years	99	17.5%	36	22.5%	95	19.0%	15	16.7%
25-34 years	334	58.9%	104	65.0%	305	61.1%	58	64.4%
>35 years	128	22.6%	20	12.5%	86	17.2%	17	18.9%
Missing	6	1.1%	0	0.0%	13	2.6%	0	0.0%
Ethnic background								
Dutch	204	36.0%	52	32.5%	161	32.3%	33	36.7%
Non Dutch	315	55.6%	92	57.5%	272	54.5%	56	62.2%
Missing	48	8.5%	16	10.0%	66	13.2%	1	1.1%
Neighbourhood								
Privileged neighbourhood	294	51.9%	80	50.0%	246	49.3%	49	54.4%
Underprivileged neighbourhood	273	48.1%	80	50.0%	253	50.7%	41	45.6%
Time entering second stage of labour								
0-8 hr	194	34.2%	36	22.5%	177	35.5%	35	38.9%
8-18 hr	240	42.3%	79	49.4%	194	38.9%	20	22.2%
18-24 hr	113	19.9%	27	16.9%	97	19.4%	31	34.4%
Missing	20	3.5%	18	11.3%	31	6.2%	4	4.4%

After the introduction of Nitrous Oxide analgesia (January - December 2011)							
D First choice Nitrous Oxide				E First choice pethidine/ EDA		Additional analgesic use	Substitution of pethidine/ EDA
D2 NO + pethidine/EDA		D3 Nitrous Oxide ALL				Column A vs. C	Column B vs. E
n=n	16 (2%)	n=n	106 (15%)	n=n	105 (15%)		
	%		%		%	p-value	p-value
						0.084	.018
12	75.0%	67	63.2%	92	87.6%		
4	25.0%	39	36.8%	13	12.4%		
						0.106	.162
7	43.8%	22	20.8%	32	30.5%		
9	56.3%	67	63.2%	57	54.3%		
0	0.0%	17	16.0%	14	13.3%		
0	0.0%	0	0.0%	2	1.9%		
						0.547	.356
5	31.3%	38	35.8%	37	35.2%		
10	62.5%	66	62.3%	59	56.2%		
1	6.3%	2	1.9%	9	8.6%		
						0.390	.048
8	50.0%	57	53.8%	40	38.1%		
8	50.0%	49	46.2%	65	61.9%		
						0.730	.135
7	43.8%	42	39.6%	34	32.4%		
1	6.3%	21	19.8%	33	31.4%		
8	50.0%	39	36.8%	15	14.3%		
0	0.0%	4	3.8%	23	21.9%		

for the period before and after the introduction of N_2O analgesia. For these periods three subgroups were made; receiving no analgesia, first choice receiving pethidine and/or EDA, first choice receiving N_2O. The latest subgroup was divided into two groups, namely those who received N_2O only and those who received N_2O and pethidine and/or EDA.

We compared women in 2010 receiving no analgesia with women in 2011 receiving no analgesia (column A vs. column C). Women in 2011 were less likely to be primiparous, <19 or >35 years, and living in a privileged neighborhood, compared to women who received no analgesia in 2010, suggesting these women received extra analgesia (N_2O). However, no statistical differences were observed.

Comparing women receiving pethidine and/or EDA in 2010 with women receiving pethidine and/or EDA in 2011 (column B vs. column E) shows that women receiving pethidine and/or EDA in 2011 were less likely to be multiparous, aged between 25-34, to live in a privileged neighborhood, and entering second stage of labor between 8-18 hour, compared to woman receiving pethidine and/or EDA in 2010, suggesting that a substitution of pethidine and/or EDA analgesia by N_2O analgesia occurs within these women. Statistical differences were only observed for parity and neighborhood.

After starting N_2O analgesia, 16 (15%) out of 106 women needed additional analgesia (column D2). Compared to women who only used N_2O analgesia, women who needed additional analgesia (column D1 vs. column D2), were more likely to be primiparous, younger, and to live in a privileged neighborhood (see table 8.2). Those women had the same profile as the group who did not use analgesia prior to the introduction of N_2O analgesia (primiparous women, privileged neighborhood).

Table 8.3 shows characteristics for women who only used N_2O analgesia and who used N_2O and additional analgesia. Physiology measurements, such as blood pressure and pulse, did not differ between the two groups and did not change over time. The Ramsay score of sedation did not differ between the two groups but increases with time (15-30 min interval; mean 2.2, range 1.0 – 4.0). However, never reaches a level compatible with total loss of consciousness. The duration of delivery was shorter ($p<0.05$) in the group only using N_2O analgesia compared to those needing additional analgesia. Fetal events including suspicion of fetal distress, the presence of low Apgar score (<7 after 5 min) and hospital admission of the child seem more present in the group who needed additional analgesia. Outcomes on physiology measurements, such as blood pressure and pulse, did not differ between the two groups and were within psychological limits. Statistical differences cannot be calculated because of low numbers.

Table 8.3 Maternal and fetal outcome subdivided for woman who used Nitrous Oxide only and women who used Nitrous Oxide and additional analgesia (pethidine/EDA)

	Nitrous Oxide only			Nitrous Oxide and pethidine/EDA		
	n = 90			n = 16		
Physiology Measurements	mean	range		mean	range	
Systolic Bloodpressure (mmHg)	121	85	166	130	96	163
Diastolic Bloodpressure (mmHg)	74	52	140	71	53	87
Heart rate (beats/min)	86	60	113	81	55	100
Ramsay score of sedation	1.2	1.0	3.0	2.1	1.0	4.0
Cervical dilatation (cm)	5	3	6	5	3	9
Duration delivery (minutes)* ****	862	104	5,919	1,945	420	6,039
Duration Nitrous Oxide (minutes)	79	10	300	89	10	302
Occurrence of complications	n	%		n	%	
Nausea	26	28.9%		3	18.8%	
Vomiting	9	10.0%		1	6.3%	
Lower state of conscious	3	3.3%		2	12.5%	
Headache	0	0.0%		0	0.0%	
Fever	0	0.0%		0	0.0%	
Fetal events	n	%		n	%	
Suspicion of fetal distress**	4	4.4%		0	0.0%	
Instrumental delivery***	16	17.8%		6	37.5%	
Low Apgar score (<7 after 5 min)	1	1.1%		0	0.0%	
Hospital admission of the child	2	2.2%		3	18.8%	

*defined as time from start contractions untill time entering second stage.
**defined as interventions because of fetal distress; fetal blood sampling during labour, episiotomy, fundus expression, instrumental delivery.
***vacuum or forceps extraction / section caesarean.
****p<0.05.

Table 8.4 shows the analgesic outcome measurements, comparing midwives and woman's rating in both the N_2O only group, and the group who needed additional analgesia. VAS scores decreased during labor. The VAS score difference between baseline and after 15-30 minutes was 1.65 (sd 2.58) for the N_2O only group compared to 1.00 (sd 1.34) for the N_2O and additional analgesia group. As stated in the methods we judge this as clinically significant for the N_2O only group.[17-19] This clinical relevant difference was observed in 70% (50/70) of

Table 8.4 Effect measurements at baseline, 15-30 min and 60 min after starting nitrous oxide analgesia subdivided for woman who used Nitrous Oxide only and women who used Nitrous Oxide and additional analgesia (pethidine/EDA)

	Nitrous Oxide only					
	baseline n=90		15-30 min n=70		60 min n=49	
Measurements during labour	mean (sd)		mean (sd)		mean (sd)	
VAS* (mean)	8.62	1.26	6.92	2.28	6.10	
Ramsay score of sedation	1.20	0.54	2.15	0.86	2.27	0.91
	%	total n	%	total n	%	total n
VAS* (%>8)	88.7%	71	47.6%	63	29.0%	31
Midwives point of view						
Pain (%very/extreme painful)	100.0%	78	82.9%	70	75.0%	36
Exhaustion (%very/extreme exhausted)	76.0%	75	52.2%	67	62.9%	35
Anxiety (%very/extreme anxious)	61.3%	75	20.0%	70	27.0%	37
Retrospective measurements						
Selfreported maternal experience						
Pain (%very/extreme painful)	86.5%	74	39.4%	66	43.8%	48
Exhaustion (%very/extreme exhausted)	15.1%	73	4.4%	68	12.2%	49
Anxiety (%very/extreme anxious)	58.1%	74	18.6%	70	18.8%	48
Selfreported maternal satisfaction						
% unsatisfied	15.9%	69				

* Visual Analoge Scale (o=no pain, 10= worst pain ever).
** Difference calculated only for patients where both observations (baseline and 15-30 min) were measured.

the N_2O only group and in 58% (7/12) of the crossover group. Scores on pain, exhaustion and anxiety showed similar trends over time, as well as differences calculated between 15-30min and baseline. The proportion of women judging their labor as very/extreme painful decreased by 32% after 15-30 minutes; judging their labor as very/extreme exhausted decreased by 11%; and judging their labor as very/extreme anxious decreased by 30%. A similar trend was observed from the midwives point of view. Within the N_2O only group 11 (12%) women were less satisfied with their pain relief compared to 3 (19%) of women in the N_2O and pethidine and/or EDA group. In retrospect women tend to rate their pain, exhaustion and anxiety after 24 hours lower than their attending midwives.

		Nitrous Oxide and EDA/Pethidine							
Difference** 15-30 min vs. baseline		baseline n=16		15-30 min n=12		60 min n=9		Difference** 15-30 min vs. baseline	
mean diff (sd)		mean (sd)		mean (sd)		mean (sd)		mean diff (sd)	
1.65	2.58	7.71	2.23	6.50	2.07	6.44	2.46	1.00	1.34
0.91	0.86	2.08	2.75	2.30	0.95	3.00	0.71	0.67	1.66
		%	total n	%	total n	%	total n		
37.5%		71.4%	14	41.7%	12	33.3%	9	27.3%	
20.3%		100.0%	15	91.7%	12	71.4%	7	9.1%	
20.3%		78.6%	14	72.7%	11	80.0%	5	0.0%	
34.4%		80.0%	15	16.7%	12	28.6%	7	81.8%	
32.8%		90.0%	10	50.0%	10	44.4%	9	36.4%	
10.9%		22.2%	9	0.0%	9	0.0%	8	9.1%	
29.7%		60.0%	10	44.4%	9	37.5%	8	9.1%	
		33.3%	9						

DISCUSSION

While analgesic options were restricted to pethidine and EDA in an obstetric unit, midwife led practice changed after the introduction of N_2O analgesia in the birthing centre. The introduction of N_2O analgesia led a substitution of pethidine and EDA analgesia in normal aged multiparous women, and to an additional increase in analgesia use predominantly within primiparous older women. Both women in whom pethidine and/or EDA was substituted with N_2O analgesia and women who previously did not use analgesia, but now used N_2O analgesia, lived more often in a privileged neighborhood.

Among those starting with N_2O analgesia, 15% needed additional analgesia. Those women had the same profile as the group who did not use analgesia prior to the introduction of N_2O analgesia (older primiparous women, privileged neighborhood). The N_2O only group showed substantial decreased rates of pain, exhaustion and anxiety during labor. Measurements by midwives during labor paralleled retrospective response 24 hours later by the mother on the same time intervals. Similar trends on the use of N_2O analgesia were observed within women who needed additional analgesia, though decreases of pain, exhaustion and anxiety were less. No major complications due to N_2O administration occurred during the study period.

To our knowledge no previous studies showed results on the changes of patient flows and analgesic use after introducing N_2O analgesia. Reports on the effectiveness of N_2O analgesia are contrasting. In an observational study Ranta et al showed no significant effect on pain relief of N_2O analgesia. However, they included women with complicated pregnancies and women whose delivery was induced, a condition known to lead to more analgesia.[20,21] Other studies showed N_2O analgesia to be an effective form of pain relief.[22,23] Unequivocal results were observed in randomized controlled trials, comparing N_2O analgesia with other analgesia or oxygen to be effective [12,17,19,24-29], while other studies show no analgesic effect of N_2O analgesia.[10,11,18,30] Most of these studies suffered from lack of power [11,17,19,24-29], lack of blinding investigators [12,18,23-30], presence of crossover effects (assuming pain during labor being consistent over time) [10,11,17-19,29,30], and presence of memory effects.[12,31]

A recent review concluded that evidence is insufficient to determine the effectiveness of N_2O for the management of labor pain compared with other, nonepidural labor pain management methods for similar reasons described above.[3] Our study results suggest N_2O analgesia to be a valid option, as women were free to move onto pethidine or EDA (e.g. no financial barrier) and did so in only 15% (16/106).

Reported rates of complications of N_2O administration were low. We observed 27% nausea and vomiting while the reported range is 5%-36%[3,17,19,23-27,32] unconsciousness was not observed, while published range is 0.4%-1%.[32-34]

Substitution of previous existing analgesic options was most prominent in multiparous women, and in this group crossover was rare (1%). Rosen et al. already concluded N_2O analgesia to be more effective among multiparous women.[8] We can only speculate on the background of this parity effect. Harrison et al suggested that N_2O analgesia was most likely to be sufficient in women whose labor is short enough to allow them to cope with pain.[22] Labor for primiparous women is characteristically longer and more painful than for multiparous.[6] Remarkably, in our study, women demanding additional analgesia started their initial N_2O treatment with an equal cervical dilatation as those for whom N_2O analgesia was

sufficient (matched for parity). N_2O analgesia was more successful in women from privileged neighborhoods. Three possible explanations include that these women may have an increased technical control, other psychological resources to cope with pain, or may have a lower incidence of other risk factors and co morbidities, leading to higher analgesic effects.[35] Data did not permit further analysis of psychological, medical, or non medical factors.

If we adopt the subjective criteria suggested by Yeo et al for a clinical relevant difference in the VAS score, we observed clinically significant differences in VAS scores after 15-30 minutes in 70% (50/70) of the N_2O only group and in 58% (7/12) of the crossover group.[29]

A strength of this study was the assessment of the implementation of N_2O analgesia on its substitution and added use by different women (patient flows) using a close historic comparison. While the observational design does not yield strong evidence on effectiveness, the naturalistic approach and almost complete cohort data allow for a fair assessment of the pros and cons of N_2O analgesia introduction.

Another strength is the use of two sources for outcome measurements, i.e. the midwife and the mother. We assessed outcome measurements using four independent techniques namely, VAS scores, 5-likert scales on pain, exhaustion and anxiety, and patient satisfaction. Patient satisfaction showed similar results as is reported in the literature (88%)[12,26,27] and was undertaken since questions were raised how to measure N_2O analgesia effectiveness.[3] Although the VAS scores seem to be sensitive to smaller changes in analgesic effect, we added the other measurements to measure outcomes from a broader perspective. Perhaps the use of VAS scores does not provide sufficiently encompassing data to measure the benefit of N_2O analgesia during labor. Nitrous Oxide may make patients feel better without reducing objective pain measurement.[3,8] 5-Likert scales were chosen since they have the advantage of supplying reference labels for grading, hereby limiting inter-individual error.[36,37]

Some study limitations merit discussion. Within our study no comparison with placebo was made. We did not apply a formal comparative study design. However, randomized controlled trials can be challenged by selective participation, since women want to receive pain medication when needed and women might have a strong preference for a certain type of analgesia.

An observational design implies that patient selection for N_2O analgesia could influence outcomes on effectiveness (selection by indication). At this stage no explicit criteria were available to guide the choice between N_2O analgesia, pethidine or EDA when analgesia was considered. As we currently do not know in whom relative performance of N_2O analgesia is best, selection effects are likely to be small. The choice for analgesia is actually based on a dual decision, combining the clinical judgment of the midwife and the preference of the woman.

Within the Dutch system additional pain medication may only be administered by the gynecologist and anesthesiologist. The increase in the use of N_2O analgesia may in part be caused by the prevention of referral and by N_2O analgesia use being the 'new' modality in midwifery led practice. Furthermore, the availability of N_2O analgesia may have led to a selection of women planning their birth in the birth centre, who want to avoid epidural or pethidine analgesia (and referral).

Although we systematically recorded the presence of complications, conclusions on the safety of nitrous oxide cannot be drawn from our study, because of low numbers.

Since these measurements include outcomes within an implementation phase, we can only speculate on the use of N_2O analgesia in the future. Increased experience in the use of N_2O by the community midwives may lead to an increase in good timing and good administration of N_2O analgesia, resulting in better effective pain relief. In addition, development of criteria to select women for whom N_2O analgesia is suitable might also help to increase N_2O analgesic effectiveness.

From this single centre study we conclude that N_2O analgesia is worthwhile to be considered as an analgesic option within birth centers adjacent to the hospital or similar settings.

REFERENCES

1. PRN. Perinatale Zorg in Nederland 2004. Utrecht: Stichting Perinatale Registratie Nederland. 2006.
2. van der Vyver M, Halpern S, Joseph G. Patient-controlled epidural analgesia versus continuous infusion for labour analgesia: a meta-analysis. Br J Anaesth. 2002 Sep;89(3):459-65.
3. Likis FE, Andrews JA, Collins MR, et al. Nitrous Oxide for the Management of Labor Pain. Comparative Effectiveness Review No. 67. (Prepared by the Vanderbilt Evidence-based Practice Center under Contract No. 290- 2007-10065-I.) AHRQ Publication No. 12-EHC071-EF. Rockville, MD: Agency for Healthcare Research and Quality. www.effectivehealthcare.ahrq.gov/reports/final.cfm. 2012.
4. Rooks JP. Nitrous oxide for pain in labor—why not in the United States? Birth. 2007 Mar;34(1):3-5.
5. Clyburn P. The use of Entonox for labour pain should be abandoned. Int J Obstet Anesth 2001 Jan;10(1):27-9.
6. Lowe NK. The nature of labor pain. Am J Obstet Gynecol. 2002 May;186(5 Suppl Nature):S16-24.
7. Bishop JT. Administration of nitrous oxide in labor: expanding the options for women. J Midwifery Womens Health. 2007 May-Jun;52(3):308-9.
8. Rosen MA. Nitrous oxide for relief of labor pain: a systematic review. Am J Obstet Gynecol. 2002 May;186(5 Suppl Nature):S110-26.
9. van der Kooy J, De Graaf JP, Kolder ZM, et al. A newly developed scavenging system for administration of nitrous oxide during labour: safe occupational use. Acta Anaesthesiol Scand. 2012 Aug;56(7):920-5.
10. Carstoniu J, Levytam S, Norman P, Daley D, Katz J, Sandler AN. Nitrous oxide in early labor. Safety and analgesic efficacy assessed by a double-blind, placebo-controlled study. Anesthesiology. 1994 Jan;80(1):30-5.

11. Volmanen P, Akural E, Raudaskoski T, Ohtonen P, Alahuhta S. Comparison of remifentanil and nitrous oxide in labour analgesia. Acta Anaesthesiol Scand. 2005 Apr;49(4):453-8.

12. Holdcroft A, Morgan M. An assessment of the analgesic effect in labour of pethidine and 50 per cent nitrous oxide in oxygen (Entonox). J Obstet Gynaecol Br Commonw. 1974 Aug;81(8):603-7.

13. CVZ. Cv. List of Obstetric Indications (in Dutch: Verloskundig Vademecum) 2003. Diemen (http://www.knov.nl/docs/uploads/Verloskundig_Vademecum_2003.pdf). 2003.

14. J VDK, JP DEG, Kolder ZM, et al. A newly developed scavenging system for administration of nitrous oxide during labour: safe occupational use. Acta Anaesthesiol Scand. 2012 Mar 7.

15. De Jonghe B, Cook D, Appere-De-Vecchi C, Guyatt G, Meade M, Outin H. Using and understanding sedation scoring systems: a systematic review. Intensive Care Med. 2000 Mar;26(3):275-85.

16. Verweij AvdL, F. Droomers, M. Rijksinstituut voor Volksgezondheid en Milieu (RIVM), Gezondheid in veertig krachtwijken. http://www.rivm.nl/bibliotheek/rapporten/270951001.pdf. 2008.

17. McGuinness C, Rosen M. Enflurane as an analgesic in labour. Anaesthesia. 1984 Jan;39(1):24-6.

18. Wee MY, Hasan MA, Thomas TA. Isoflurane in labour. Anaesthesia. 1993 May;48(5):369-72.

19. McLeod DD, Ramayya GP, Tunstall ME. Self-administered isoflurane in labour. A comparative study with Entonox. Anaesthesia. 1985 May;40(5):424-6.

20. Ranta P, Jouppila P, Spalding M, Kangas-Saarela T, Hollmen A, Jouppila R. Parturients' assessment of water blocks, pethidine, nitrous oxide, paracervical and epidural blocks in labour. Int J Obstet Anesth. 1994 Oct;3(4):193-8.

21. Guerra GV, Cecatti JG, Souza JP, et al. Elective induction versus spontaneous labour in Latin America. Bull World Health Organ. 2011 Sep 1;89(9):657-65.

22. Harrison RF, Shore M, Woods T, Mathews G, Gardiner J, Unwin A. A comparative study of transcutaneous electrical nerve stimulation (TENS), entonox, pethidine + promazine and lumbar epidural for pain relief in labor. Acta Obstet Gynecol Scand. 1987;66(1):9-14.

23. Rosen M, Mushin WW, Jones PL, Jones EV. Field trial of methoxyflurane, nitrous oxide, and trichloroethylene as obstetric analgesics. Br Med J. 1969 Aug 2;3(5665):263-7.

24. Jones PL, Rosen M, Mushin WW, Jones EV. Methoxyflurane and nitrous oxide as obstetric analgesics. II. A comparison by self-administered intermittent inhalation. Br Med J. 1969 Aug 2;3(5665):259-62.

25. Jones PL, Rosen M, Mushin WW, Jones EV. Methoxyflurane and nitrous oxide as obstetric analgesics. I. A comparison by continuous administration. Br Med J. 1969 Aug 2;3(5665):255-9.

26. Bergsjo P, Lindbaek E. Comparison between nitrous oxide and methoxyflurane for obstetrical analgesia. Acta Obstet Gynecol Scand. 1971;50(3):285-90.

27. Stefani SJ, Hughes SC, Schnider SM, et al. Neonatal neurobehavioral effects of inhalation analgesia for vaginal delivery. Anesthesiology. 1982 May;56(5):351-5.

28. Abboud TK, Swart F, Zhu J, Donovan MM, Peres Da Silva E, Yakal K. Desflurane analgesia for vaginal delivery. Acta Anaesthesiol Scand. 1995 Feb;39(2):259-61.

29. Yeo ST, Holdcroft A, Yentis SM, Stewart A, Bassett P. Analgesia with sevoflurane during labour: ii. Sevoflurane compared with Entonox for labour analgesia. Br J Anaesth. 2007 Jan;98(1):110-5.

30. Westling F, Milsom I, Zetterstrom H, Ekstrom-Jodal B. Effects of nitrous oxide/oxygen inhalation on the maternal circulation during vaginal delivery. Acta Anaesthesiol Scand. 1992 Feb;36(2):175-81.

31. Norvell KT, Gaston-Johansson F, Fridh G. Remembrance of labor pain: how valid are retrospective pain measurements? Pain. 1987 Oct;31(1):77-86.

32. Rooks JP. Safety and risks of nitrous oxide labor analgesia: a review. J Midwifery Womens Health. Nov-Dec;56(6):557-65.

33. Clinical trials of different concentrations of oxygen and nitrous oxide for obstetric analgesia. Report to the Medical Research Council of the Committee on Nitrous Oxide and Oxygen Analgesia in Midwifery. Br Med J. 1970 Mar 21;1(5698):709-13.

34. McAneny TM, Doughty AG. Self-Administered Nitrous-Oxide/Oxygen Analgesia in Obstetrics. With Particular Reference to the 'Lucy Baldwin' Machine. Anaesthesia. 1963 Oct;18:488-97.

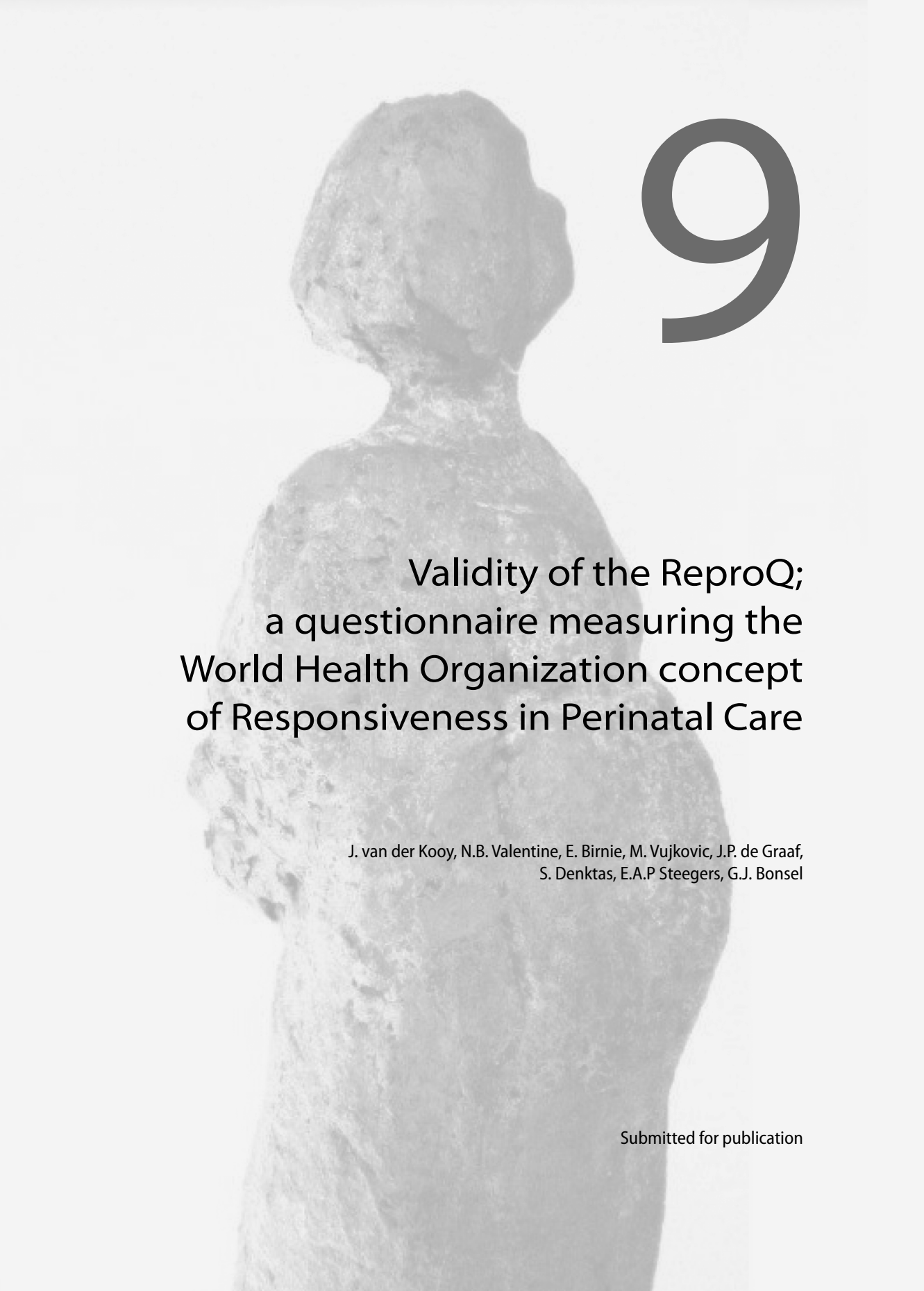
35. Waud BE, Waud DR. Calculated kinetics of distribution of nitrous oxide and methoxyflurane during intermittent administration in obstetrics. Anesthesiology. 1970 Apr;32(4):306-16.

36. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. Pain. 1975 Sep;1(3):277-99.

37. Melzack R. The short-form McGill Pain Questionnaire. Pain. 1987 Aug;30(2):191-7.

PART III

How can the performance of the Dutch perinatal care be evaluated on non clinical aspects of quality of care?



9

Validity of the ReproQ; a questionnaire measuring the World Health Organization concept of Responsiveness in Perinatal Care

J. van der Kooy, N.B. Valentine, E. Birnie, M. Vujkovic, J.P. de Graaf, S. Denktas, E.A.P Steegers, G.J. Bonsel

Submitted for publication

ABSTRACT

Background The concept of responsiveness, introduced by the World Health Organization, aims to address service quality regardless of provider, country or health system. Responsiveness is defined as aspects of the way individuals are treated and the environment in which they are treated during health system interactions.

Objectives To assess the validity of a newly developed ReproQ responsiveness questionnaire dedicated to maternal experiences during perinatal care using the eight-domain WHO concept.

Method The ReproQ was developed between October 2009 and February 2010 by translating the WHO Questionnaire items contextually to perinatal care. Psychometric properties (feasibility, construct validity including Cronbach's alpha and item correlations per domain, discriminative validity) were empirically assessed in a cohort of Dutch women, two weeks post partum.

Results A total of 171 women consented to participation. Feasibility: the interviews lasted between 20 and 40 minutes and the overall item missing rate was 8%. Construct validity: mean Cronbach's alphas for the antenatal, delivery and postpartum phase were: 0.73 (range 0.57-0.82), 0.84 (range 0.66-0.92), and 0.87 (range 0.62-0.95) respectively. The item-own domain scale correlations within all phases were considerably higher than most of the item-other domain scale correlations. Within the antenatal, delivery and post partum phase, the eight factors explained 69%, 69%, and 76% of variance respectively. Discriminative validity: overall responsiveness mean sum scores were higher for women whose children were not admitted for pediatric care.

Conclusion The ReproQ questionnaire demonstrated satisfactory psychometric properties to describe the quality of perinatal care in the Netherlands, with the potential to discriminate between quality of care levels.

INTRODUCTION

The concept of responsiveness, introduced by the World Health Organization (WHO) in 2000, offers health services the opportunity to capture patient's experiences of their interactions on eight predefined domains.[1-2] In recent years, the Dutch system of perinatal care has been the subject of discussion with reference to the particularly high national perinatal mortality rates compared to other countries in Europe.[3]

The performance of perinatal care is often judged by its endpoints such as perinatal mortality and morbidity and costs. However, quality of care literature supports the view that non-clinical aspects of health care such as service quality are important too, and indirectly affect clinical outcomes. [4-6] Better service quality is thought to increase compliance with medical treatment and to improve the transfer of information and appropriate utilization of health services.[7-10] Governments of Western countries increasingly acknowledge the importance of the non-clinical aspects of quality of care and incorporate these when the provision of care is monitored.[11-12]. Within The Netherlands, this study is the first to apply an international concept to the Dutch perinatal care system.[13-15] This paper is also among the first global attempts to apply the international concept of responsiveness to a particular part of the health system – the perinatal system.

The Dutch perinatal health care system can be regarded as a chain of health care provisions, which chronologically covers antenatal care, delivery and post partum care. Antenatal, delivery-related and post partum care are provided by different caregivers with different responsibilities, for different risk groups, and in different settings. Independently operating community midwives provide care for low-risk pregnant women (primary healthcare) while gynecologists providing in-hospital care for high-risk women (secondary and tertiary care). Most women receive post partum care by a community midwife.

The concept of responsiveness seems apt to evaluating non-clinical aspects of this system as its domains cover the relevant aspects of the patient-provider interaction along the chain of health care. WHO defines responsiveness along eight domains as the way an individual is treated and the environment in which she is treated during health system interactions and additionally enables quantitative trade-offs between service quality and clinical outcome when these vary across systems.[2] It contains non-financial, non-clinical predefined domains that reflect respect for human dignity and interpersonal aspects of the care process. The independent position of these responsiveness domains is supported by human rights law.[2,16-17] Responsiveness built on the early work of the Consumer Assessment of Health Plans Study (CAHPS) through an inter-agency collaboration (1999-2003), bringing

it to an international stage.[18] It tries to capture the patient's real experience by referring to an actual event in contrast to patient satisfaction questionnaires. Literature has shown that expectations may strongly influence patient satisfaction, which makes international comparison challenging since these are influenced by economic and political influences.[19-22]

Adopting this concept, the ReproQ questionnaire was developed, dedicated to maternal experiences during perinatal care. It followed closely the WHO concept with contextual adjustment to perinatal care. The aim of this study is to investigate the psychometric properties of the ReproQ.

METHODS

Questionnaire

The WHO developed a survey, which was administrated between 2000-2001 under the auspices of the Multi-Country Survey Study on Health and Health Systems Responsiveness (MCS Study) and again in 2002-2003 under the World Health Survey (WHS).[1-2] (http://www.who.int/responsiveness/surveys/en/) The ReproQ was developed between October 2009 and February 2010, and its questions were derived from these WHO questionnaires.

The ReproQ questionnaire was developed to assess the responsiveness of perinatal health care system in the Netherlands and is based on the same eight domains identified in WHO's review of the patient satisfaction and quality of care literature, i.e. Dignity, Autonomy, Confidentiality, Communication, Prompt Attention, Social Consideration (labeled initially as *Access to Social Support* or *Access to Family and Community Support*), Quality of Basic Amenities, and Choice and Continuity.

These domains are claimed to be of universal importance in all health systems, during any client-system interaction (including personal and non-personal health services) and for the population's interaction with insurers and other administrative arms of the health system. While it is recognized that persons may differ regarding the relative importance of each domain, and that specific domains may be of extra relevance in particular health care interactions, it is assumed that the quality of any interaction is sufficiently covered by these eight domains.[2]

The ReproQ asks the same questions for the three different phases of perinatal care: antenatal phase (the period from the onset of pregnancy until the onset of delivery), delivery phase and post partum phase (covering the first ten days after childbirth).

Rather than pointing to a single event, or the last visit (as in MCS and WHS), we prefered women's judgement on all antental visits rather than just a single visit, as the latter is easily biased by a particular incident. A similar argument applies to the postnatal maternity care. Delivery, however, seems appropriate for the conventional event-like approach. Within this framework the setting and professional items where contextually translated to the perinatal care system (e.g. 'doctor' was translated into 'midwife' or 'gynecologist'). If two different health care professionals could be involved (e.g. 'midwife' and '(maternity) nurse' during delivery), similar questions within each domain were repeated for each health care professional separately.

Each phase was covered by the above mentioned eight domains, with 2-7 items per domain. The standardized response mode consisted of 5 options: 'very good', 'good', 'moderate', 'bad', and 'very bad'. The ReproQ consisted of 104 questions on responsiveness (25 antenatal, 40 delivery, 39 postpartum phase) and 29 questions for maternal and health care characteristics.

Questions from the WHO questionnaire were translated into Dutch according to a predefined protocol. First, questionnaires were translated by the research team. Expert meetings consisting of gynecologists, midwives, nurses, public health experts and researchers were held to judge the translation and comprehensiveness of the item list. Many among these professionals had working experience in English speaking countries. Next, backward translation of each question was then performed and comparison was made with the original English questionnaire. Improvements were made and final consensus was reached on each question.

The completeness of domains was judged in terms of being comprehensive (are all non-clinical areas covered, which patients and professionals put forward either as positive experience or negatively as complaint, and would it cover organizational changes of the system), and in terms of being balanced (have all domains included about equal importance). For each domain the candidate pool of items was checked whether each item fitted to the domain definition sufficiently. As this could differ per phase, this was discussed for each phase separately (e.g. the item 'quality of the food' during antenatal visits was excluded). Finally we asked the experts to check whether all the domains would remain valid under ongoing and anticipated organizational changes in perinatal care. All stakeholders agreed on the final list that the stated requirements were met.

Six primiparous and multiparous pregnant women were invited to judge the feasibility of the draft version of the questionnaire. Figure 9.1 shows the eight domains and items for the antenatal phase.

Dignity	Were physical examinations and treatments done in a way that respected your privacy?
	Did the examination rooms ensure your privacy?
	Were you treated with respect by your health care provider?
Autonomy	How well were you involved in making decisions regarding your examinations or treatments?
	Were you able to refuse examinations or treatments?
	Were you asked permission before testing or starting treatment?
Confidentiality of Information	Were consultations carried out in a manner that protected your confidentiality?
	Was confidentiality kept on the information provided by you?
	Was your medical record kept confidential?
Communication	How well were things explained by your health care provider in a way you could understand?
	Was written information provided in such a way you could understand?
	Were you encouraged to ask questions about your health problems, treatment and care?
	Were you given time to ask questions about your health problem or treatment?
	Was information on the health service's contact, location and parking information clear to you?
Prompt Attention	How well did you receive prompt attention at your health service?
	How did you experience the waiting time after you asked for help?
	How well was the accessibility by phone?
	How do you rate the travel time to your health service?
Social Consideration	Did the health care provider facilitate the support of your relatives and friends?
	Was the home situation taken into consideration when planning an appointment?
Quality of basic amenities	How do you rate the quality of the hygiene of the toilets?
	How do you rate the overall quality of the surroundings, for example, space, seating, fresh air and cleanness?
	How do you rate the quality of the food?
Choice and Continuity of Health Care Provider	Were you able to choose your own health care provider?
	Were you able to use other health care services other than the one you usually went to?
	How well was the continuity of care by one health care provider?
	Were you able to choose your own place of delivery?

Figure 9.1 The eight domains with the items given for the antenatal phase.

Study population and data collection

Study approval was granted by the Medical Ethical Committee, Erasmus Medical Centre, Rotterdam, the Netherlands, no MEC2012207.

To investigate the psychometric properties of responsiveness questions for each phase, women were recruited from three midwifery practices in Rotterdam, the Netherlands, between February 2010 and March 2011 (almost all women, regardless their health utility, receive post partum care by a community midwife in the Netherlands). Women or their partners were required to speak and understand Dutch sufficiently. Written informed consent was obtained.

At the post partum visit two weeks after delivery, a subset of women was invited by their own midwife for study participation which implied a 30 minute face-to-face interview with an independent interviewer. This interview was held at another site, usually at home. The face-to-face interviews were carried by ten trained independent interviewers (medical students) and covered questions on maternal and health characteristics, and on responsiveness outcomes on the antenatal, delivery and post partum phase. Interviewees were invited to respond to all questions, yet never forced to.

Data handling and analysis

Records were regarded 'missing at the record level' if all scores of all phases were missing (antenatal, delivery and post partum phase). If women had responded partially, the responses were evaluated per phase.

If all the items of one phase were missing, this record was excluded from the analysis of that phase. This implies that occasionally respondents were excluded from one phase while they were included in the analysis of other phases.

We investigated the responsiveness questions' psychometric properties stratified for the antenatal phase (the period from the onset of pregnancy until the onset of delivery), the delivery phase and postpartum phase (covering the first ten days after delivery).

Sumscores

Unweighted sumscores per domain were calculated and transformed into 1-10 domain scores to enhance comparability among domains with different numbers of items. Transformation was done as score = 1 + 9* ([sumscore – lowest sum possible)] / [largest sum possible – lowest sum possible]). E.g., a domain that contains 3 items each with a 5-point response mode, displays a possible score range from 3 to 15. The transformed sumscore would then be 1 + 9* ([sumscore - 3] / [15 – 3]). If sumscore in an individual would be 11, her transformed score would be 1 + 9*([11-3]/[15-3])= 1 + 9* (8/12) = 7. This transformation procedure was repeated for each domain in each phase separately.

Psychometric tests

The following psychometric properties of the ReproQ were evaluated: feasibility, reliability and validity (discriminant and construct validity).

Feasibility was expressed as rates of missing items per domain. The literature provides little indication of acceptable survey response rates or inappropriate non-response rates but, in general, missing item rates below 20% can be considered acceptable.[1] Furthermore, we assessed systematic missing rates per domain by age, education, race, communication and health utilization (missing rates significantly higher among defined subgroups).

Scale scores were described by in terms of scale mean, SD, range, floor and ceiling effects, and percentiles.

Reliability was assessed as scale internal consistency and item correlations. Scale internal consistency was assessed using Crohnbach's alpha. Amidst varying standards in the literature, we considered 0.70 to be an acceptable alpha coefficient.[23] Average inter-item, average item-own scale and average item-other scale correlation were assessed with standardized correlation coefficients, with acceptable correlations defined as Pearson's correlation coefficients (r) > 0.40.[24] We expected higher average inter-item and average inter-own scale correlations compared to average inter-other scale correlations.

Discriminative validity was assessed by comparing subgroups expected to differ in terms of responsiveness. It was hypothesized that women whose child was not admitted to the hospital for pediatric care would report better outcomes than women whose child was hospitalized. Differences in overall mean sum scores (adding all domains) were calculated and tested with Student t-tests per phase.

Construct validity was assessed as the domain structure of factor loadings obtained with confirmatory factor analysis using maximum likelihood factor analysis with oblique promax rotation, extracting eight (fixed) factors.

RESULTS

A total of 274 women were invited for study participation of whom 94 declined. Reasons for non-participation included the time burden, feeling at unease having a stranger visit their home, and logistic reasons such as incorrect phone number, or incorrect address. 180 women (66%) agreed to be interviewed. Of these seven interviews (7/180, 4%) were cancelled and two interviews (2/180, 1%) had to be stopped because the respondents did

not speak sufficiently Dutch and no translator could be made available. The remaining 171 interviews were used for analysis. The interviews took between 20 and 40 minutes. Table 9.1 describes the baseline characteristics of the participants.

Table 9.1 Maternal characteristics and outcome

Variable	n	%
Maternal Age* **		
<19 years	3	2%
20-25 years	15	9%
25-34 years	119	70%
>35 years	33	19%
missing	1	1%
Parity*		
Primiparous	64	37%
Multiparous	44	26%
missing	97	57%
Ethnic background*		
Dutch	97	57%
Non Dutch	74	43%
Education*		
Low	6	4%
Middle	75	44%
High	90	53%
Marital status*		
single	141	82%
relationship/married	30	18%
Neighbourhood		
privileged neighbourhood	84	49%
underprivileged neighbourhood	87	51%
Proficiency (speaking) Dutch*		
good/excellent	152	89%
weak/poor	18	11%
missing	1	1%
Care process*		
start antenatal care with midwife, not referred	61	36%
start antenatal care with midwife, referred during antenatal phase to gynaecologist	37	22%

(*continued*)

Table 9.1 Continued

Variable	n	%
start antenatal care with midwife, referred during birth phase	57	33%
start antenatal care with gynaecologist	16	9%
Hospital admission of child*		
No admission	145	85%
Admission	26	15%

* p-value <0.05 (t-test).
**mean age 31 (95%CI 30.0-31.7).

The mean maternal age was 31 years (95%CI 30.3–31.7). The majority of mothers were primiparous (57%). Of the 171 mothers 74 (43%) were of non Dutch origin, 18 (11%) spoke weak/poor Dutch as judged by the interviewer and almost 50% came from an underprivileged neighborhood. 81 (48%) had a low/middle education and 30 (18%) were single mothers. Referral to gynecologists occurred in approximately 55% of women. Post partum hospital admission occurred in 26 (15%) of all newborns.

Table 9.2 describes the missing rates per domain, for each phase separately. The results of four women with all domain scores in the delivery care phase missing were excluded.

The item missing rate over all phases was 8% (1,349 out of 17,624 questions). Missing rates per domain were all below the predefined threshold of 20%. Missing rates were highest in the delivery phase (8%) and the domains; 'Social Consideration' (delivery phase) and 'Autonomy' (post partum phase).

Systematic missing rates were significantly more present among women of Dutch origin within the domain 'Confidentiality' and pertained mainly to the delivery phase. There were no systematic missing by age, educational level and health utilization.

Table 9.3 displays the transformed scores per domain and phase (1-10 scale). Mean transformed scores were relatively high (7.1–8.4) as were the median scores (7.2-7.8). Floor effects were observed in 0.6% of women. Mean scores and ceiling effects differed most across the domains in the antenatal phase and least across the domains in the post partum phase. Comparisons of the domain scores across the three phases showed a non-uniform pattern, suggesting that respondents judged each phase separately.

The Cronbach's alpha averaged over the domains was 0.73 (range 0.57-0.82) for the antenatal phase, 0.84 (range 0.66-0.92) for the delivery care phase and 0.87 (range 0.62-0.95) for the post partum phase. For all phases the domain 'Quality of Basic Amenities' had lowest alphas.

Table 9.2 Missing values given per domain after excluding records with total missings

	Antenatal Care		(N=171)	Birth care		(N=167)	Post Partum Care		(N=171)	Total		(N=509)
Domain	Total Items	missing(n)	%	Total Items	n	%	Total Items	n	%	Total Items	n	%
Dignity	513	7	1%	835	21	3%	855	55	6%	2,203	83	4%
Autonomy	513	57	11%	501	57	11%	855	112	13%	1,869	226	12%
Confidentiality	513	28	5%	1,002	88	9%	1,026	91	9%	2,541	207	8%
Communication	855	35	4%	1,002	49	5%	1,026	65	6%	2,883	149	5%
Prompt Attention	684	28	4%	1,169	111	9%	684	58	8%	2,537	197	8%
Social Consideration	342	16	5%	501	65	13%	855	60	7%	1,698	141	8%
Quality of Basic Amenities	342	8	2%	501	45	9%	513	35	7%	1,356	88	6%
Choice and Continuity	513	45	9%	1,169	130	11%	855	83	10%	2,537	258	10%
Total	4,275	224	5%	6,680	566	8%	6,669	559	8%	17,624	1,349	8%

Table 9.3 Mean(SD), Range, percentage Floor and Ceiling, Crohnbach's α

Domain	No. of items	Mean	SD	Range		% Floor	% Ceiling	25th %tile	50th %tile	75th %tile	Crohnbach's α
Dignity											
Antepartum Phase	3	8.4	1.1	5.5	10.0	0.0%	21.6%	7.8	7.8	9.3	0.73
Birth Phase	5	8.1	1.1	1.1	10.0	0.0%	11.7%	7.8	7.8	9.1	0.86
Post Partum Phase	5	7.9	1.3	3.3	10.0	0.0%	12.3%	7.8	7.8	8.2	0.87
Autonomy											
Antepartum Phase	3	7.8	1.2	3.3	10.0	0.0%	8.2%	7.0	7.8	8.5	0.73
Birth Phase	3	7.7	1.4	1.4	10.0	0.0%	8.8%	7.7	7.8	7.8	0.87
Post Partum Phase	5	7.5	1.7	1.9	10.0	0.0%	0.6%	7.3	7.8	7.8	0.94
Confidentiality											
Antepartum Phase	3	8.0	1.1	4.0	10.0	0.0%	14.0%	7.8	7.8	8.5	0.82
Birth Phase	6	7.8	1.4	1.4	10.0	0.0%	12.3%	7.8	7.8	7.8	0.78
Post Partum Phase	6	7.7	1.4	1.8	10.0	0.0%	13.5%	7.4	7.8	7.8	0.94
Communication											
Antepartum Phase	5	7.7	1.2	3.3	10.0	0.0%	5.3%	7.3	7.8	8.2	0.80
Birth Phase	6	7.8	1.3	1.3	10.0	0.0%	9.9%	7.4	7.8	8.1	0.92
Post Partum Phase	6	7.6	1.7	1.0	10.0	0.6%	11.7%	7.4	7.8	8.1	0.95
Prompt Attention											
Antepartum Phase	4	7.1	1.4	1.0	10.0	0.6%	2.3%	6.6	7.2	7.8	0.67
Birth Phase	7	7.7	1.3	1.3	10.0	0.0%	7.0%	7.1	7.8	8.4	0.83
Post Partum Phase	4	7.7	1.7	1.0	10.0	0.6%	12.9%	7.2	7.8	8.9	0.89
Social Consideration											
Antepartum Phase	2	7.1	1.8	1.0	10.0	0.6%	8.2%	5.5	7.8	7.8	0.76
Birth Phase	3	7.6	1.6	1.6	10.0	0.6%	11.1%	7.0	7.8	7.8	0.87
Post Partum Phase	5	7.8	1.4	3.3	10.0	0.0%	8.2%	7.3	7.8	8.7	0.84

Quality of Basic Amenities											
Antepartum Phase	2	7.5	1.4	3.3	10.0	0.0%	10.5%	6.6	7.8	7.8	0.57
Birth Phase	3	7.6	1.4	1.4	10.0	0.0%	8.2%	7.0	7.8	8.5	0.66
Post Partum Phase	3	7.4	1.5	1.8	10.0	0.0%	6.4%	7.0	7.8	7.8	0.62
Choice and Continuity											
Antepartum Phase	3	7.3	1.7	1.0	10.0	0.6%	7.0%	6.3	7.8	7.8	0.77
Birth Phase	7	7.2	1.5	1.5	10.0	0.0%	5.3%	6.5	7.6	7.8	0.88
Post Partum Phase	5	7.1	1.7	1.0	10.0	0.6%	7.0%	6.4	7.8	7.8	0.89

Table 9.4 describes the average item-scale correlations. The average inter-item correlations were 0.49 for the antenatal phase, 0.58 for the delivery phase and 0.63 for the post partum phase. Average inter-item correlations were relatively low for the domains 'Prompt Attention' and 'Quality of Basic Amenities'. The item-own scale correlations for each phase separately were considerably higher than most of the corresponding item-other scale correlations. The overall average item-own scale correlation was 0.56 for the antenatal phase, 0.68 for the delivery phase and 0.73 for the post partum phase.

Mean overall sum scores were higher for women whose child was not admitted after delivery: 61.8 (sd 7.4) versus 58.3 (sd 5.1) (p=0.02) in the antenatal phase; 61.9 (sd 8.4) versus 57.9 (sd 7.7) (p=0.06) in the delivery phase; and 62.1 (sd 9.2) versus 55.2 (sd 13.0) (p=0.01) in the post partum phase.

The fixed eight factors explained 69% of the variance in the antenatal phase, 69% in the delivery phase and 76% in the post partum phase. Table 9.5 shows the final result of oblique promax rotated factor loadings of the delivery phase (the patterns of the antenatal and post partum phase were about similar). Items that were expected to belong to one domain are outlined. The rotated solution of grouped items generally confirmed the hypothesized domain taxonomy within the delivery and post partum phase. For the antepartum phase however, the hypothesized domain taxonomy was less present with regard to 'Social Consideration' and 'Choice and Continuity', which appeared to be associated with other domains.

Table 9.4 Average inter-item, item-own scale, item-other scale, correlations of Responsiveness Domains (n=109)

	Antenatal Care (N=168)			Birth care (N=160)			Post Partum Care (N=150)		
Domain	Average inter-item correlation	Average item-own scale correlation*	Average item-other scale correlation	Average inter-item correlation	Average item-own scale correlation*	Average item-other scale correlation	Average inter-item correlation	Average item-own scale correlation*	Average item-other scale correlation
Dignity	0.48	0.56	0.37	0.56	0.69	0.56	0.58	0.70	0.63
Autonomy	0.46	0.55	0.36	0.68	0.74	0.57	0.77	0.85	0.64
Confidentiality	0.63	0.69	0.37	0.65	0.70	0.57	0.73	0.82	0.63
Communication	0.44	0.58	0.35	0.67	0.78	0.56	0.76	0.85	0.64
Prompt Attention	0.34	0.46	0.36	0.41	0.58	0.54	0.67	0.76	0.63
Social Consideration	0.61	0.61	0.37	0.70	0.76	0.57	0.51	0.65	0.63
Quality of Basic Amenities	0.40	0.40	0.37	0.42	0.49	0.56	0.36	0.44	0.62
Choice and Continuity	0.53	0.62	0.36	0.53	0.68	0.54	0.62	0.73	0.62

* each item was correlated with the applicable scale excluding the item under consideration from the scale score.

Table 9.5 Promax rotated factor solution for the Birth phase

	Factor								
Factor Name	Confidentiality	Choice	Dignity	Prompt attention	Autonomy	Communication	Quality of Basic Amenities		Unique variance
Respect shown during examinations (midwife)	.032	.096	.670	-.004	-.053	-.004	-.178	.243	.305
Examination room suitable to provide privacy	.090	.018	.792	.018	-.163	-.221	.137	-.013	.277
Treated with respect (midwife)	-.103	-.138	.871	.037	-.059	.001	.057	.071	.310
Respect shown during examinations (nurse)	-.004	.031	.768	-.051	.092	.193	-.095	-.061	.267
Treated with respect (nurse)	-.056	-.103	.568	.032	.017	.244	.067	.029	.265
Involved in making a decision regarding your examinations or treatments	-.163	.041	.065	-.058	.895	-.051	.010	.080	.279
Able to refuse examinations or treatments	.011	-.029	-.162	.044	1.009	.059	-.102	-.006	.242
Asked permission before testing or starting treatment	.059	.187	-.033	-.046	.693	-.067	-.046	.003	.377
Protecting your confidentiality during consultations (midwife)	.635	.055	.135	-.003	.062	-.048	.100	.039	.234
Confidentiality kept on provided information (midwife)	.773	.076	.099	-.098	.055	.014	.078	-.064	.179

(*continued*)

Table 9.5 Continued

	Factor								
Factor Name	Confidentiality	Choice	Dignity	Prompt attention	Autonomy	Communication	Quality of Basic Amenities		Unique variance
Confidentiality of patients' medical records preserved (midwife)	.843	.063	.086	-.117	.065	-.020	-.070	.033	.161
Protecting your confidentiality during consultations (nurse)	.548	-.081	-.051	.091	-.192	-.126	.071	.156	.584
Confidentiality kept on provided information (nurse)	.955	-.079	-.066	.144	-.108	.113	-.023	-.050	.156
Confidentiality of patients' medical records preserved (nurse)	.872	.012	-.160	.099	.046	.089	.003	.016	.140
Information clearly explained (midwife)	.236	-.221	-.072	.164	.225	.236	.000	.359	.281
Information about other treatment options (midwife)	.065	.057	-.022	-.155	.157	.181	.093	.641	.211
Encouraged to ask questions about diseases, treatment and care (midwife)	.105	.010	.131	.006	-.068	.201	-.059	.694	.233
Information clearly explained (nurse)	.032	-.139	-.041	.013	.057	.720	.085	.188	.234
Information about other treatment options (nurse)	.009	.079	-.062	-.078	.079	.701	.085	.225	.172

Encouraged to ask questions about diseases, treatment and care (nurse)	.065	.047	.124	-.036	-.090	.699	-.045	.227	.216
Experience of the waiting time when arriving on the place of delivery	.154	.031	-.102	.743	-.121	.012	-.051	.005	.423
Experience of the waiting time on examinations	-.108	.109	.030	.683	.003	-.001	-.096	.399	.257
Experience of the waiting time after you asked for help (midwife)	-.056	.196	.072	.724	-.068	-.101	-.031	-.014	.360
Accessibility by phone (midwife)	-.026	.013	.188	.082	.089	.099	.270	.080	.493
Travelling time to the place of birth	.082	-.144	.085	.514	.215	.038	.028	-.174	.463
Experience of the waiting time after you asked for help (nurse)	.012	.034	.030	.406	-.023	.118	.234	-.002	.387
Accessibility by phone (nurse)	.055	.164	-.020	.195	-.033	.237	.218	-.023	.402
Facilitate the support of relatives and friends (midwife)	.144	.025	.000	-.064	-.098	-.032	.937	-.017	.253
Consideration of home situation when planning appointments/ examinations	.123	.101	.089	-.121	.190	.018	.466	.119	.253
Facilitate the support of relatives and friends (nurse)	-.091	.075	.050	.077	-.057	.219	.690	-.026	.287

(*continued*)

Table 9.5 Continued

Factor Name	Factor							Unique variance
	Confidentiality	Choice	Dignity	Prompt attention	Autonomy	Communication	Quality of Basic Amenities	
Hygiene of the toilets and examination rooms.	-.068	-.086	.244	.167	.206	.108	.151	-.160	.503
Comfort of the examination rooms and waiting rooms	.157	-.002	.256	.303	.201	-.108	.005	-.241	.440
Quality of the food	-.078	.237	.143	.171	.193	-.153	-.039	-.028	.667
Able to choose own health care provider (midwife)	.119	.611	-.116	.124	.032	-.264	.144	.132	.383
Able to be referred to a medical specialist (midwife)	.267	.398	.127	-.165	.069	-.062	.014	.074	.404
Presence of different health care providers (midwife)	.062	.622	-.039	.056	.066	.128	-.096	.043	.295
Continuity of care by one health care provider (midwife)	.059	.434	.148	.179	.034	-.003	-.094	.119	.321
Able to choose own health care provider (nurse)	-.207	.683	-.123	.020	.077	-.062	.344	-.016	.327
Presence of different health care providers (nurse)	-.043	.760	-.002	-.029	-.076	.495	-.169	-.150	.269
Continuity of care of one health care provider (nurse)	-.045	.492	-.024	.049	-.101	.558	.065	-.131	.294
Total variance explained; 69%									

DISCUSSION

With the support of both patients and caregivers we developed and translated the WHO's concept of responsiveness into the ReproQ instrument to measure responsiveness in the Dutch obstetric care system antenatally, during delivery and post partum. ReproQ appeared to be a potential instrument for reporting perinatal service quality from the patient's perspective. The original domain structure proved to be comprehensive, as judged by the stakeholders, while the presumed structure of items was grossly confirmed after data analysis. The ReproQ demonstrated satisfactory psychometric properties to describe the quality of perinatal care in the Netherlands, with the potential to discriminate between quality of care levels.

Particular strengths are discussed below. Firstly, the eight domains were chosen based on a pre-existent philosophical structure, as identified in WHO's review of the patient satisfaction and quality of care literature, which also included the examination of different survey instruments[16] Secondly, the independent value of the domains are supported by human rights law which argues that the responsiveness features of a health system are important in their own right.[2,16-17] Thirdly, in contrast to patient satisfaction questionnaires, responsiveness tries to capture the patient's real experience, since literature has shown that expectations strongly influence patient satisfaction. Expectations may be influenced by economic influences, political influences, prior experiences and socio-demographic characteristics.[19-22] Fourthly, responsiveness aimed to develop a universal concept (e.g. developing and developed countries, different ethnicities, different care systems, etc.)[2]

The Responsiveness concept is challenged by a number of issues. Firstly, although responsiveness aims to measure the patient's actual experience, it is still disturbed by at least some extent of 'subjectivity'. Secondly, the concept of Responsiveness does not include financial barriers since these are evaluated separately by the WHO. Thirdly, capturing responsiveness by a limited number of questions with fixed answering categories is quite challenging. Combining qualitative research and different (quantitative) survey techniques, one can produce a richer, more valid, and more reliable findings than when adopting qualitative or quantitative methods alone.[25]

Participation rates were equal or higher than the participation rates found in other perinatal satisfaction studies.[26-28] Participation rates were equal to participation rates found in surveys measuring similar domains of quality of care[1,18,29] , and better than obtained by WHO's Multi-Country Survey (MCS) study administered in the Netherlands

in 2001 (59%). Comparisons are made with the MCS Study that was conducted in the Netherlands in 2001 as the questionnaire contained multiple items for each responsiveness domain, whereas the subsequent World Health Survey only contained one question per domain.[1-2]

The domain missing rates were below 20%, which according to the literature can be considered acceptable.[1] However, within the framework of the MCS study, slightly lower overall missing rate was reported (5.0% instead of 8.0% in our case).[30] However, our survey contained three phases whereas the MCS study focused on a single interaction in the previous 12 months, and was shorter (on average 25 minutes).[28,31-32] As found in the MCS, the domain missing rate was highest for the domains of 'Autonomy' and 'Choice and Continuity' which are typically cognitively demanding domains. Across phases of perinatal care, the missing rate was highest for delivery phase. Most likely this is the consequence of some items pointing to service events that do not always taking place.

The scale properties were satisfactory. A floor effect was almost absent as is frequently the case in positive-skewed assessments of self-reported health or self-rated experiences of (maternity) care. [26-28,33] There was surprisingly less skewing towards use of the most positive category (ceiling effects) compared to other surveys e.g. in the MCS. [28,31-32] Importantly for health equity reasons, missing rates were not selective for different age groups, educational level or health care utilization. For unknown reasons selective missing was found among Dutch women on the domain 'Confidentiality' during the delivery phase.

The average inter-item correlation of the items used to construct the domains were high. They were higher on average than recorded for the face-to-face survey in the general population conducted by International Research Associates (INRA EUROPE) for WHO as part of the MCS Study in 2001 (0.6 across domains and phases by ReproQ, versus 0.3 in MCS).[1] Within each phase and for all domains, the questionnaire's internal consistency was good. Crohnbach's Alpha coefficients in ReproQ were similar compared to the CAHPS and WHO surveys [1,18], except for the domain 'Quality of Basic Amenities', which showed poor alphas in all phases. This domain contained questions about sanitary hygiene, comfort of waiting room and quality of food. It can be argued that these elements of basic amenities were too diverse to achieve internal consistency (see Table 9.5).

Overall, the taxonomy of domains in the ReproQ study was fairly well preserved, across all phases, although a weaker result in the antenatal phase was present. Additionally, observed

results in the antenatal phase showed lower inter-item correlations relative to results for other phases. This could possibly be due to factors such as; recall bias introduced by assessing all phases together, contamination by pregnancy outcome, focusing on one particular event or the heterogeneity in measurements since antenatal care consists of multiple visits. Underlying patterns are still to be explored. One may consider presenting a questionnaire on the antenatal phase within the antenatal phase, separately from a questionnaire on the delivery and post partum phase.

The ability of the instrument to discriminate between good and less good experiences will be of paramount importance for its future use. We found some promising test results. The respondents clearly expressed different opinions on their experiences in the different phases of perinatal care. Discrimination between women whose infants were subsequently admitted to hospital was reflected in the lower sumscores across all phases. However, to test the difference in mean responsiveness of the delivery phase between the mothers whose infant was hospitalised and the mothers whose infants were not hospitalised (mean difference: 3.8, pooled SD 6.5), at least 194 mothers had to be included in the analysis (type I error = 0.05 (two-sided), power = 0.80, control/case-ratio: 6/1).

Different responses on antenatal sumscores may reflect a true outcome on non clinical aspect of care or may be a contamination by pregnancy outcomes. This again stresses presenting a questionnaire on the antenatal phase separately.

Test-retest reliability was not performed in this stage. Reasons were the burden for the participants and the main difficulty to avoid memory effects of this demanding interview.

Overall, the ReproQ questionnaire, which was derived from the WHO concept of responsiveness, overall demonstrated satisfactory psychometric properties to describe the non-clinical aspects of perinatal care in a three-site study in the Netherlands. The instrument has the potential to discriminate quality of care across the different phases and for different levels of experiences. In general, psychometric properties were in line with results obtained for other survey instruments that have been tested and promoted as part of quality assessment effort. In conclusion, given the lack of comparable instruments and the overall favorable study results, we feel that this unique adaptation of the WHO responsiveness questionnaire to evaluate the various phases of maternity care has been relatively successful. With some minor adaptations we believe that this questionnaire can be used to evaluate the quality of perinatal care in The Netherlands and beyond.

REFERENCES

1. Valentine NB, Bonsel GJ, Murray CJ. Measuring quality of health care from the user's perspective in 41 countries: psychometric properties of WHO's questions on health systems responsiveness. Qual Life Res. 2007 Sep;16(7):1107-25.
2. Valentine NB, de Silva A, Kawabata K, Darby C, Murray CJL, Evans DB (eds). Health system responsiveness: concepts, domains and measurement. Geneva: World Health Organisation; 2003.
3. Ravelli AC, Tromp M, van Huis M, et al. Decreasing perinatal mortality in The Netherlands, 2000-2006: a record linkage study. J Epidemiol Community Health. 2009 Sep;63(9):761-5.
4. Campbell SM, Roland MO, Buetow SA. Defining quality of care. Soc Sci Med. 2000 Dec;51(11): 1611-25.
5. Sitzia J, Wood N. Patient satisfaction: a review of issues and concepts. Soc Sci Med. 1997 Dec;45(12):1829-43.
6. Wensing M, Jung HP, Mainz J, Olesen F, Grol R. A systematic review of the literature on patient priorities for general practice care. Part 1: Description of the research domain. Soc Sci Med. 1998 Nov;47(10):1573-88.
7. McPake B. User charges for health services in developing countries: a review of the economic literature. Soc Sci Med. 1993 Jun;36(11):1397-405.
8. Ware JE, snyder MK, Wright WR, Davies AR (eds). Defining and measuring patient satisfaction with medical care. Evaluation and Program Planning. Volume 6: 247-263; 1983.
9. Wouters A. Essential national health research in developing countries: heath care financing and the quality of care. Internation Journal of Health Planning and Management. 1991;6: 253-271.
10. Gilson L, Alilio M, Heggenhougen K. Community satisfaction with primary health care services: an evaluation undertaken in the Morogoro region of Tanzania. Soc Sci Med. 1994 Sep;39(6):767-80.
11. Wensing M, Elwyn G. Methods for incorporating patients' views in health care. BMJ. 2003 Apr 19;326(7394):877-9.
12. Greener I. Towards a history of choice in UK health policy. Sociol Health Illn. 2009 Apr;31(3):309-24.
13. Christiaens W, Gouwy A, Bracke P. Does a referral from home to hospital affect satisfaction with childbirth? A cross-national comparison. BMC Health Serv Res. 2007;7:109.
14. Wiegers TA. The quality of maternity care services as experienced by women in the Netherlands. BMC Pregnancy Childbirth. 2009;9:18.
15. Wiegers TA, van der Zee J, Keirse MJ. Transfer from home to hospital: what is its effect on the experience of childbirth? Birth. 1998 Mar;25(1): 19-24.
16. de Silva A (ed). A framework for measuring responsiveness. Global Programme on Evidence for Health Policy Discussion Paper. No. 32. Geneva: World Health Organisation.
17. Gostin L (ed). The domains of health responsiveness: a human rights assessment. Geneva: World Health Organisation; 2002.
18. Hargraves JL, Hays RD, Cleary PD. Psychometric properties of the Consumer Assessment of Health Plans Study (CAHPS) 2.0 adult core survey. Health Serv Res. 2003 Dec;38(6 Pt 1):1509-27.
19. Blendon RJ, Kim M, Benson JM. The public versus the World Health Organization on health system performance. Health Aff (Millwood). 2001 May- Jun;20(3):10-20.
20. Hall JA, Irish JT, Roter DL, Ehrlich CM, Miller LH. Satisfaction, gender, and communication in medical visits. Med Care. 1994 Dec;32(12):1216-31.
21. Murphy-Cullen CL, Larsen LC. Interaction between the socio-demographic variables of physicians and their patients: its impact upon patient satisfaction. Soc Sci Med. 1984;19(2):163-6.
22. Murray CJ, Kawabata K, Valentine N. People's experience versus people's expectations. Health Aff (Millwood). 2001 May-Jun;20(3):21-4.
23. Nunnally JCB, I.H. Psychometric theory. (3rd ed) New York: McGraw-Hill. 1994.
24. Gandek B, Ware JE, Jr., Aaronson NK, et al. Tests of data quality, scaling assumptions, and reliability of the SF-36 in eleven countries: results from the IQOLA Project. International Quality of Life Assessment. J Clin Epidemiol. 1998 Nov;51(11):1149-58.
25. Foundation NS. User friendly Handbook for mixed evaluations. 1997.

26. Smith LF. Development of a multidimensional labour satisfaction questionnaire: dimensions, validity, and internal reliability. Qual Health Care. 2001 Mar;10(1):17-22.

27. Smith LF. Postnatal care: development of a psychometric multidimensional satisfaction questionnaire (the WOMBPNSQ) to assess women's views. Br J Gen Pract. 2011 Oct;61(591):e628-37.

28. Redshaw M, Heikkila K. Delivered with care: a national survey of women's experience of maternity care 2010. NPEU https://www.npeu.ox.ac.uk/delivered-with-care. 2010.

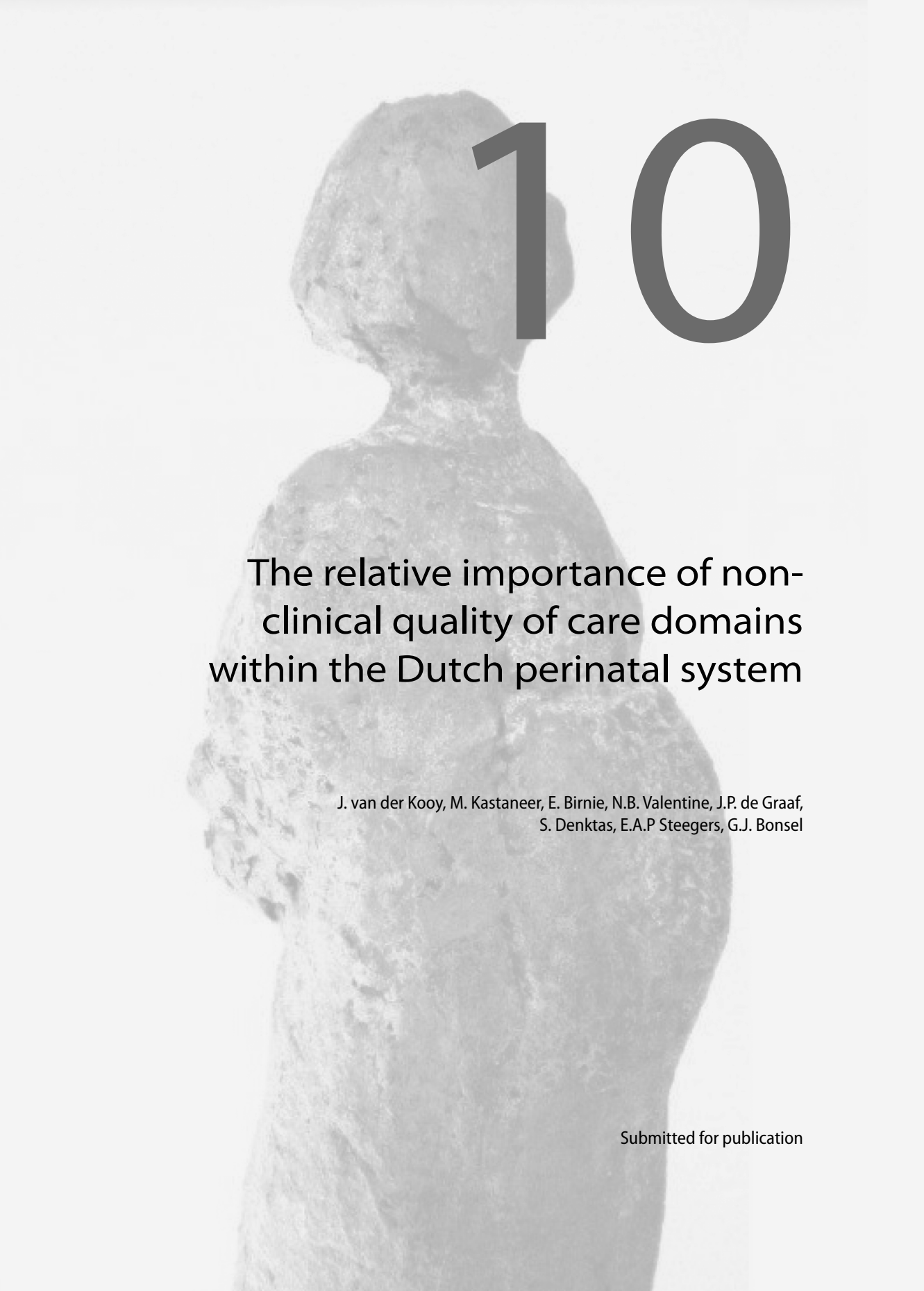
29. Jenkinson C, Coulter A, Bruster S. The Picker Patient Experience Questionnaire: development and validation using data from in-patient surveys in five countries. Int J Qual Health Care. 2002 Oct;14(5):353-8.

30. T. Bedirhan Ustün SC, Maria Villanueva, Lydia Bendib, Can Çelik, Ritu Sadana, Nicole Valentine, Juan Ortiz, Ajay Tandon, Joshua Salomon, Yang Cao, Xie Wan Jun, Emre Ozaltin, Colin Mathers and Christopher J.L. Murray. WHO Multi-country Survey Study on Health and Responsiveness 2000-2001 GPE Discussion Paper 37. http://wwwwhoint/responsiveness/papers/gpediscpaper37pdf. 2001.

31. Vujkovic M, Birnie E, Bonsel GJ. Maternal postnatal care utilisation in Rotterdam 2008-2009. The Rotterdam Public Health Authority Report (GGD Rotterdam Rijnmond) (in press). 2011.

32. Demott K, Bick D, Norman R, et al. Clinical Guidelines And Evidence. Review For Post Natal Care: Routine Post Natal Care Of Recently Delivered Women And Their Babies http://www.nice.org.uk/nicemedia/live/10988/30146/30146.pdf. 2006.

33. Peterson WE, Charles C, DiCenso A, Sword W. The Newcastle Satisfaction with Nursing Scales: a valid measure of maternal satisfaction with inpatient postpartum nursing care. J Adv Nurs. 2005 Dec;52(6):672-81.



10

The relative importance of non-clinical quality of care domains within the Dutch perinatal system

J. van der Kooy, M. Kastaneer, E. Birnie, N.B. Valentine, J.P. de Graaf, S. Denktas, E.A.P Steegers, G.J. Bonsel

Submitted for publication

ABSTRACT

Background The concept of responsiveness was introduced by the World Health Organization (WHO) to address service quality in an international comparable way. Responsiveness is defined as aspects of the way individuals are treated and the environment in which they are treated during health system interactions. The concept consists of eight domains. These domains are not of equal importance to everyone since the importance of domains may be influenced by individual factors and country factors.

Objectives We developed a responsiveness questionnaire for the Dutch perinatal system based on the general WHO format and we investigated whether the assigned domain weights in this specific context were equivalent, particularly in predefined target groups.

Method Women were recruited from three midwifery practices in Rotterdam, The Netherlands. Responsiveness was measured and additionally two assignments were added, focusing on the ranking of the eight domains in terms of perceived importance and ranking the items within a domain.

Results For the majority of respondents, women who recently gave birth, Communication and Dignity were the two most important domains and Choice and Continuity and Social Consideration were the two least important; however individual heterogeneity is considerable as is shown by low overall concordance. Individual factors had a modest effect on the ranking order and hardly affected the ranking position of the two most and two least important domains.

Conclusion This study shows that individual factors do not have a significant effect on the relative importance of domains. Therefore we conclude that all perinatal subpopulations can use the same set of equally valued quality ('responsiveness') domains. Herewith judgements on caregivers can be made across heterogeneous subpopulations.

INTRODUCTION

The concept of responsiveness, as introduced by the World Health Organization (WHO) in 2000, is one of the few available approaches to address service quality in an internationally comparable way. WHO defines responsiveness as the way an individual is treated and the client orientation of the environment in which he or she is treated during health system interactions.[1] The WHO declares responsiveness together with health outcome and financial fairness as the primary intrinsic goals of health system performance.[2] Human rights law supports the view that the responsiveness features of a health system are important and valuable in their own right.[1,3-4] In the WHO approach, responsiveness contains the non-financial, non-clinical qualities of care, conveniently categorized in eight domains that are assumed to be of equal importance.[2]

These eight domains actually are not of equal importance to every individual. The importance may be influenced by specific individual factors, type of care (e.g. inpatient, outpatient, emergency care) and country factors in general. For universal application of one responsiveness scale, it is essential to demonstrate the absence of substantial heterogeneity of domain weights across individuals, subpopulations or countries.

Valentine et al. compared the importance of these eight domains across countries and across different subpopulations within countries using responsiveness data from the Multi-Country Survey Study on Health and Health Systems Responsiveness (MCS Study). As expected, convergence in rankings was stronger across subpopulations within countries than across countries, but both these effects were modest.[5] Nevertheless, responsiveness domains are ranked variously among different target groups and healthcare services.[6-8] When comparing different healthcare institutions within the Dutch perinatal system, one might consider using different weights for the responsiveness domains for different subpopulations.

We developed a responsiveness questionnaire for the Dutch perinatal system based on the general WHO format and we investigated whether the assigned domain weights in this specific context were equivalent, particularly in predefined target groups.

METHODS

Study settings and samples

A cross-sectional study was conducted in Rotterdam between February 2010 and March 2011, The Netherlands, where various levels of public facilities for perinatal healthcare are available, i.e. community midwifery practices, several regional hospitals, one university

hospital and one midwife-led birth centre. Almost all women receive postpartum care by a midwife (>95%). Women attended in three midwifery practices were randomly invited to participate on the study at the postpartum visit two weeks after delivery. Inclusion criteria were; women or their partners were required to speak and understand the Dutch language sufficiently; and written informed consent.

Study participation implied a 60 minute face-to-face in-depth interview with an thoroughly trained independent interviewer. The interviewers approached these women by phone and a home interview was scheduled. Interviewers were not rewarded for their efforts. An effective study sample of at least 200 women was aimed at. The study was approved by the local Medical Ethical Committee of the Erasmus Medical Centre, Rotterdam, The Netherlands.

ReproQ

Interviews were held using the ReproQ. This questionnaire was developed between October 2009 and February 2010 to assess the responsiveness of the perinatal healthcare system in the Netherlands. Its questions were derived and translated from WHO questionnaires[1,9], and were made specific for perinatal care. [unpublished data van der Kooy et al] The ReproQ consisted of eight responsiveness domains that were categorized as 'respect for persons', i.e. Dignity, Autonomy, Confidentiality, Communication, and 'client orientation', i.e. Prompt Attention, Social Consideration (labelled initially as Access to Social Support or Access to Family and Community Support), Quality of Basic Amenities, and Choice and Continuity. Each domain contained three to five items [unpublished data van der Kooy et al].

Assignments

The interview covered responsiveness outcomes (not presented in this paper) and two assignments. Women were first asked to rank the domains in terms of overall importance, and subsequently to rank the items within each domain. The ranking was forced, no ties could be made. Data collected was not shared with the caregiver.

Variables

Social characteristics were maternal age (<30 years versus >30 years), ethnic background (Dutch versus non-Dutch), education level (low versus middle/high), marital status (single versus relationship/married), neighbourhood (privileged versus underprivileged, based on 4-digit zip-codes and a public list of deprived, zip-code based, neighbourhoods issued by

the Dutch government), employment status (one or both parents employed versus both parents unemployed) and social support (≤2 persons versus ≥3 persons).

Medical and obstetric history related characteristics were parity (primiparous versus multiparous), planned pregnancy (yes versus no), obstetric history (present versus absent, based on maternal report of mother and/or child outcomes requiring intervention of a gynaecologist in a previous pregnancy) and psychiatric history (present versus absent),

Healthcare process characteristics were onset of antenatal care (<13 weeks versus ≥13 weeks' gestation) and the care process, which existed of four possible pathways: antenatal care and childbirth with a community midwife without referral to the gynaecologist; onset of antenatal care with a community midwife with referral to the gynaecologist during the antenatal phase (when risk factors arise during antenatal care); onset of antenatal care with a community midwife with referral to the gynaecologist during childbirth (when risk factors arise during childbirth); and onset of antenatal care with a gynaecologist while staying with the gynaecologist throughout the whole pregnancy.

Outcome characteristics were an adverse outcome for the child (present versus absent, based on maternal report of oxygen shortage, (possible) congenital anomaly, infection, small for gestational age, premature birth and hospital admission of the child), hospital admission of the mother after delivery (yes versus no) and intervention during delivery, i.e. caesarean section or instrumental intervention during birth (yes versus no).

Data analysis

First we calculated the mean ranking of the untransformed scores for each domain separately, by averaging the ranks assigned by each respondent. The assigned rank of the domains was given a numerical value, e.g., rank number 1 was given eight points, rank number 2 was given seven points. In addition to this untransformed rank score, we applied two weighted transformations. In transformation I, we adapted this calculation to emphasize the most important domains. Instead of assigning numerical weights of 8 (most important), 7, 6, 5, 4, 3, 2 and 1 (least important), we assigned squared weights of 8^2 (most important), 7^2, 6^2, 5^2, 4^2, 3^2, 2^2 and 1^2 (least important). This transformation acknowledges ordinality but emphasizes the most important domains. Transformation II again respects ordinality but emphasizes the extreme ranks (parabolic weights), either high or low: 9 points given to the most important domain, 4, 1, 0, 0, -1, -4, and -9 to the least important domain. As with the untransformed ranks, the average domain scores were calculated for each of the eight domains separately.

To compare ranks across subgroups we used the untransformed averages of the total group for each domain as reference. To describe the different patterns of the ranking for subpopulations, we established whether for any domain its ranking was affected or not; and if so, whether the ranking position changed one place ('simple reversal') or whether changes in ranking position involved two or more positions ('relevant change') compared to reference ranking of untransformed averages. If changes in ranking were present, this was assumed most important among the highest ranking domains.

Finally we used the non-parametric Kendall's coefficient of concordance for ranks (W) to evaluate whether the rankings of the eight domains within subpopulations were concordant ('agreement among raters'). Kendall's W ranges from 0 to 1 with the following interpretations: Kendall's W ≤ 0.3: very weak to weak within-group agreement, implying that the confidence in ranks is low; Kendall's W between 0.3 and 0.7: confidence in ranks is fair; Kendall's W ≥ 0.7: strong agreement; and Kendall's W=1.0: complete agreement among raters. As a rule of thumb, a Kendall's W ≥ 0.5 can be interpreted as a moderate agreement across and within subpopulation [10]. We argue that different weights for different subpopulations should be applied when; Kendall's W is 0.5 or higher in the total population and relevant changes were observed in subgroups compared to the total reference group, or, when Kendall's W is much higher in subgroups compared to the total reference group.

Kendall's W was first performed for all domains, we checked whether restriction to the two most important and the two least important domains changed conclusions to assess the hypothesis that extremes show more concordance for the study population and across subpopulations.

In addition, Kendall's W was determined for the items within the domains. A simple weight was used, e.g. when a domain consisted of three items the highest item in ranking was given a weight of 3 and the lowest item in ranking was given a weight of 1. Missing values were excluded from analyses. Data were analysed using SPSS software version 17.0.

RESULTS

Characteristics of respondents

A total of 274 women were invited for study participation; 180 women (66%) agreed to be interviewed. Of the 180 interviews planned, seven interviews (4%) were cancelled and two interviews (1%) were cut short because the respondents did not speak Dutch and no

translator could be made available. A total of 171 interviews (95%) were used for analysis. 125 women (73%) completed the assignment of ranking the eight domains and 157 women (91%) completed the assignment of ranking the items within the domains.

Table 10.1 Characteristics of respondents

	N	%	
ALL	171	100	PRN (%)
Social characteristics			
Maternal age			
≤30 years	73	43	44
Ethnic background			
Non-Dutch	77	45	42
Education			
Low	81	47	
Marital status			
Single	30	18	
Neighbourhood			
Underprivileged	87	51	31
Income			
Both parents unemployed	21	12	
Social support			
Maximum 0-2 persons	41	24	
Medical and obstetric history			
Parity			
Primiparous	97	57	49
Planned pregnancy			
No	47	27	
Obstetric history**			
Present	50	29	
Psychiatric history			
Present	13	8	
Healthcare process			
Start antenatal care			
≥13 weeks	24	14	
Referral process			
Start with midwife, not referred	61	36	30

(*continued*)

Table10.1 Continued

	N	%	
ALL	171	100	PRN (%)
Start with midwife, referred during antenatal phase to gynaecologist	37	22	32
Start with midwife, referred during birth phase to gynaecologist	57	33	25
Start with gynaecologist	16	9	13
Outcome			
Adverse outcome child***			
Present	75	44	
Hospital admission mother after delivery			
Yes	71	42	
Intervention during delivery****			
Yes	72	42	

*PRN (Perinatal Registration in The Netherlands): percentages from four large cities: Amsterdam, Rotterdam, The Hague, Utrecht.

**Obstetric history is based on maternal report of maternal / child outcomes requiring intervention of a gynaecologist in a previous pregnancy.

***Adverse outcome child is based on maternal report of oxygen shortage, (possible) congenital anomalie, infection, small for gestational age, premature birth and hospital admission of the child.

**** Caesarian section or instrumental delivery.

The mean age of respondents was 31.0 years (95% CI 30.3 – 31.7). 77 (45%) of respondents were of non-Dutch origin and 81 (47%) were poorly educated. Ninety-seven (57%) of the women were pregnant for the first time. Care was provided by a community midwife throughout the whole pregnancy for 61 women (36%) and provided by the gynaecologist for 16 women (9%). About half (55%) of the women was referred to the gynaecologist antenatally or during childbirth. Forty-four percent of the newborn had an adverse outcome. Compared to the averages in the four largest cities in the Netherlands, our study population shows an overrepresentation of some of these characteristics (table 1).

Importance of responsiveness domains

From figure 10.1 and table 10.2 follows the general importance of each domain expressed as the mean ranking score using three different methods of weighting. Usage of a weighting scheme that added more importance to the domain highest in rank (squared weights) or a transformation that added more importance to the domains both highest and lowest in rank (parabolic weights) did not change their positions in the overall ranking compared to

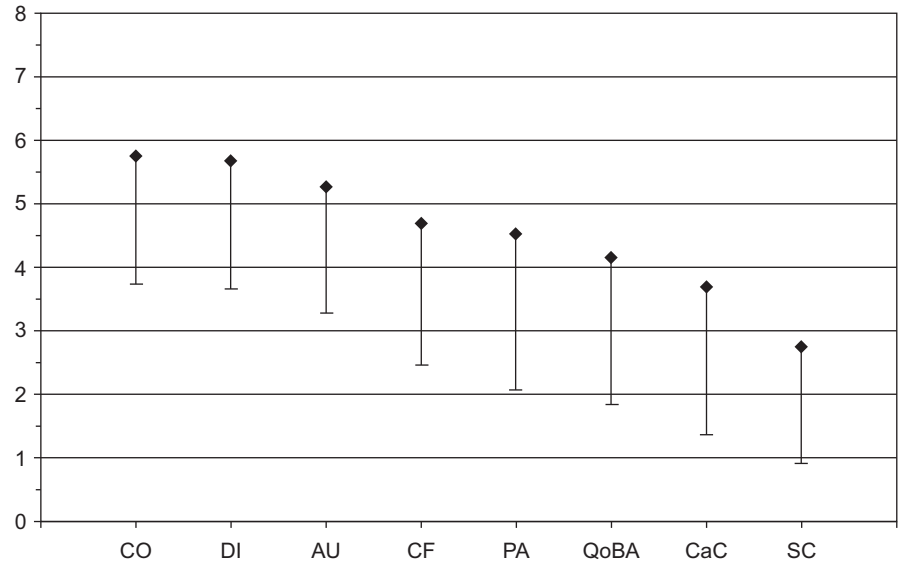


Figure 10.1 Importance of responsiveness domains according to mean raw scores.
DI = Dignity, AU = Autonomy, CF = Confiadentially, CO = Communication, PA = Prompt Attention, SC = Social Consideration, QoBA = Quality of Basic Amenities, CaC = Choice and Continuity.

the untransformed average domain scores as reference. Communication and Dignity were the two highest ranking domains across all selected characteristics. Choice and Continuity and Social Consideration were generally ranked lowest. The mean scores of the domains did not differed much, due to the overlap of the in-between ranks (3-6).

Table 10.2 Importance of responsiveness domains according to assigned rank

	simple weight*			important domains emphasized**			extreme domains emphasized***		
	mean score	(95% CI)	rank	mean score	(95% CI)	rank	mean score	(95% CI)	rank
Dignity	5.7	(5.3 – 6.0)	2	35.9	(32.3 – 39.5)	2	2.2	(1.4 – 3.0)	2
Autonomy	5.2	(4.9 – 5.5)	3	30.7	(27.4 – 34.0)	3	1.1	(0.4 – 1.8)	3
Confidentiality	4.6	(4.3 – 5.0)	4	25.7	(22.2 – 29.3)	4	0.4	(-0.4 – 1.14)	4
Communication	5.7	(5.3 – 6.0)	1	36.0	(32.2 – 39.9)	1	2.7	(1.9 – 3.5)	1
Prompt Attention	4.4	(4.0 – 4.8)	5	24.9	(21.2 – 28.6)	5	-0.3	(-1.2 – 0.6)	5
Social Consideration	2.7	(2.4 – 3.0)	8	10.6	(8.1 – 13.2)	8	-3.5	(-4.3 – -2.7)	8
Quality of Basic Amenities	4.1	(3.7 – 4.5)	6	21.5	(17.9 – 25.1)	6	-0.7	(-1.5 – 0.2)	6
Choice and Continuity	3.7	(3.3 – 4.0)	7	18.6	(15.1 – 22.0)	7	-1.8	(-2.7 – -1.0)	7

* rank untransformed: 8 (= most important), 7, 6, 5, 4, 3, 2, 1 (= least important).
** squared weight: 8^2 (= most important), 7^2, 6^2, 5^2, 4^2, 3^2, 2^2, 1^2 (= least important).
*** parabolic weight: 9 (= most important), 4, 1, 0, 0, -1, -4, -9 (= least important).

Importance of responsiveness domains for different subpopulations

Table 10.3 depicts the rankings of domains for predefined subgroups compared to the overall untransformed ranking of domains in the total reference group (Table 10.2). Table 10.3 also shows the number of domains where the ranking differed either one position (Δrank = 1) or two or more positions (Δrank≥2).

Table 10.3 Importance of person characteristics on domain rank (only changes shown)

	CO	DI	AU	CF	PA	QoBA	CaC	SC	Δ rank = 1	Δ rank ≥ 2
Overall	1	2	3	4	5	6	7	8		
Social characteristics										
Maternal age										
≤30 years (n=58)	2	1		5	4				4	0
Ethnic background										
Non-Dutch (n=57)									0	0
Education										
Low (n=64)		3	2			7	6		4	0
Marital status										
Single (n=23)			7	5	4	3	6		3	2
Neighbourhood										
Underprivileged (n=68)	2	1		6	4	5			4	1
Income										
Both parents unemployed (n=16)	2	1	5	3	6	4			4	2
Social support										
Maximum 0-2 persons (n=32)									0	0
Medical and obstetric history characteristics										
Parity										
Primiparous (n=74)	2	1							2	0
Planned pregnancy										
No (n=38)									0	0
Obstetric history*										
Present (n=39)						7	6		2	0
Psychiatric history										
Present (n=12)	3	1	4	2	6	5			4	2
Healthcare process Start antenatal care										
≥13 weeks (n=19)	2	1			6	7	5		4	1

Referral process									
Start with midwife, not referred (n=46)	2	1		5	4			4	0
Start with midwife, referred during antenatal phase to gynaecologist (n=26)		3	2		7		5	2	2
Start with midwife, referred during birth phase to gynaecologist (n=41)	2	1	5		3			2	2
Start with gynaecologist (n=12)				5	7		4	1	2
Outcome									
Adverse outcome child**									
Present (n=59)	2	1				5		4	0
Hospital admission mother after delivery									
Yes (n=55)					6	5		2	0
Intervention during delivery***									
Yes (n=52)								0	0

D I = Dignity, AU = Autonomy, CF = Confiadentially, CO = Communication, PA = Prompt Attention, SC = Social Consideration, QoBA = Quality of Basic Amenities, CaC = Choise and Continuity.
***Obstetric history is based on self reported mother or child outcomes requiring intervention of a gynaecologist in a previous pregnancy.
****Adverse outcome child is based on self reported oxygen shortage, (possible) congenital anomalie, infection, small for gestational age, premature birth and hospital admission of the child.
***** Need of intervention during birth.

In most subgroups rankings remained unchanged (25%) or differed one position (44%). The ranking of non-Dutch women, women who had a maximum of two persons for social support, women with an unplanned pregnancy and women who had an intervention during childbirth was fully identical to the overall ranking. Relevant changes i.e. rankings that differed ≥2 positions were found mostly within the social and healthcare process characteristics but not in women who experienced an adverse outcome. Relevant changes occurred frequently in the domains that were neither ranked most important nor ranked least important, with some notable exceptions. Women with a psychiatric history mentioned Confidentiality as second most important compared to 4th position in the overall ranking. Women who started antenatal care ≥13 weeks of gestation and women who were antenatally cared for by a gynaecologist attached more value to Choice and Continuity by assigning fourth or fifth ranks .

Concordance of ranks for responsiveness domains

Table 10.4 shows the concordance of ranks within the total reference group and for subgroups. Generally, concordance was very weak (Kendall's W of 0.113) and hardly improved for the subgroups. When considering only the two most important domains, Dignity and Communication, and the two least important domains, Choice and Continuity and Social Consideration, the Kendall's W increased (0.062 – 0.183) for all subgroups. Agreement was still far from strong. Both concordance analyses confirmed the earlier results which treated the ranks as numerical scores (table 10.3).

Table 10.4 Correlation of domain rank (Kendall's W) ; subgroups compared to the population

	all domains
Overall	.113
Social characteristics	
Maternal age	
≤30 years	.133
Ethnic background	
Non-Dutch	.095
Education	
Low	.096
Marital status	
Single	.100
Neighbourhood	
Underprivileged	.133
Income	
Both parents unemployed	.119
Social support	
Maximum 0-2 persons	.100
Medical and obstetric history	
Parity	
Primiparous	.119
Planned pregnancy	
No	.089
Obstetric history*	
Present	.107
Psychiatric history	
Present	.191
Healthcare process	
Start antenatal care	
≥13 weeks	.201

Referral process	
Start with midwife, not referred	.150
Start with midwife, referred during antenatal phase to gynaecologist	.339
Start with midwife, referred during birth phase to gynaecologist	.159
Start with gynaecologist	.304
Outcome	
Adverse outcome child**	
Present	.161
Hospital admission mother after delivery	
Yes	.135
Intervention during delivery***	
Yes	.176

*PRN (Perinatal Registration in The Netherlands) :percentages from four large cities: Amsterdam, Rotterdam, The Hague, Utrecht.
**Obstetric history is based on maternal report of maternal / child outcomes requiring intervention of a gynaecologist in a previous pregnancy.
***Adverse outcome child is based on maternal report of oxygen shortage, (possible) congenital anomalie, infection, small for gestational age, premature birth and hospital admission of the child.
**** Caesarian section or instrumental delivery.

Concordance of ranks for items within domains

Kendall's W was also determined for the items within the domains. Concordance was very weak for the items within most domains, indicating substantial individual heterogeneity of ranking of items. A weak to moderate agreement was found only for items within the domains Communication and Quality of Basic Amenities.

Table 10.5 Ranking the items within the domains

Domain	Overall Kendall's W
Dignity	.060
Autonomy	.192
Confidentiality	.020
Communication	.314
Prompt Attention	.106
Social Consideration	.146
Quality of Basic amenities	.262
Choice and Continuity	.055

DISCUSSION

In this study we investigated the importance of the eight responsiveness domains within the Dutch perinatal health system. For the majority of respondents Communication and Dignity were the two most important domains and Choice and Continuity and Social Consideration were the two least important; however individual heterogeneity is considerable as is shown by low overall concordance as well as in the total group and the separate subgroups. Subgroup characteristics had a modest effect on the ranking order and hardly affected the ranking position of the two most and two least important domains. These results suggest that this quality of care measurement can be used in different subpopulations without specific subgroup adaptation.

Study strengths

Firstly, this study had a high response rate. In our study, about two third of the approached women agreed to participate, while a response rate of 30% has been proposed as reasonable for patient satisfaction surveys and a response rate of 50% is considered to be quite high[11-12].

Our study covered all subgroups in Rotterdam, including groups which often refrain from participation in satisfaction surveys (people with a psychiatric history; those with a low social economic status; those with a low educational level; people without paid work and Muslim people)[13-14]. However, since only people from urban areas participated in this study, the study population is primarily representative for Dutch urban areas. Thirdly, in this study a simple ranking exercise was used. Empirical results from studies using multiple methods have not led to the emergence of a single preferred method.[15] Other techniques easily place a greater cognitive burden on respondents.[16] The ranking data allow our results to be compared to other studies.[5]

Study limits

Barriers to the generalizability are the selection of the study population, whom primarily comes from a Dutch urban area, and, the non-participation of women who did not understand the Dutch language sufficiently. If translation could be arranged, this was usually done by a family member. Since translation was not performed by a professional translator, both random and systematic bias may have been introduced. Secondly, various ethnicities joined into one category of non-Dutch resulted in a heterogeneous subpopulation. Preferences regarding the importance of domains may differ within ethnic subgroups.[17] Thirdly, no analysis was performed on non-participants; ethical permission for this study did not include non-respondents analysis.

Comparing results with other studies

We found Dignity and Communication to be the most important domains. In contrast to our study the WHO surveys [1,5,18] and Liabsuetrakul et al [7] found Prompt Attention and Dignity to be the most important domains, followed by Communication in third place. The preference for Prompt Attention may be due to the fact that in the WHO surveys and Liabsuetrakul et al[7] Prompt Attention was operationalized in terms of geographical access and access in case of emergencies. In our study Prompt Attention was focussed upon waiting times, as in our country perinatal care is available for all pregnant women because of complete insurance coverage and dense supply of perinatal services. Results from other studies which focussed upon waiting times support our results that Prompt Attention was of less importance.[17,19]

The two least important domains were Choice and Continuity and Social Consideration. Lisabsuetrakul et al[7] performed a responsiveness study within perinatal healthcare and found similar results to ours regarding the least important domains. Choice and Continuity seems to be of less importance within perinatal healthcare, because within the WHO surveys this domain was found more important. This may be due to the fact that care for pregnant women is often acute and therefore a specific health provider usually cannot be assigned, unless women strongly request for one and the caregiver agrees.

The concordance in ranking of the responsiveness domains is low, even within subpopulations. Fletcher et al[20] investigated patients' priorities for medical care and found low concordance using a method quite close to ours.[21]

Unexpected findings

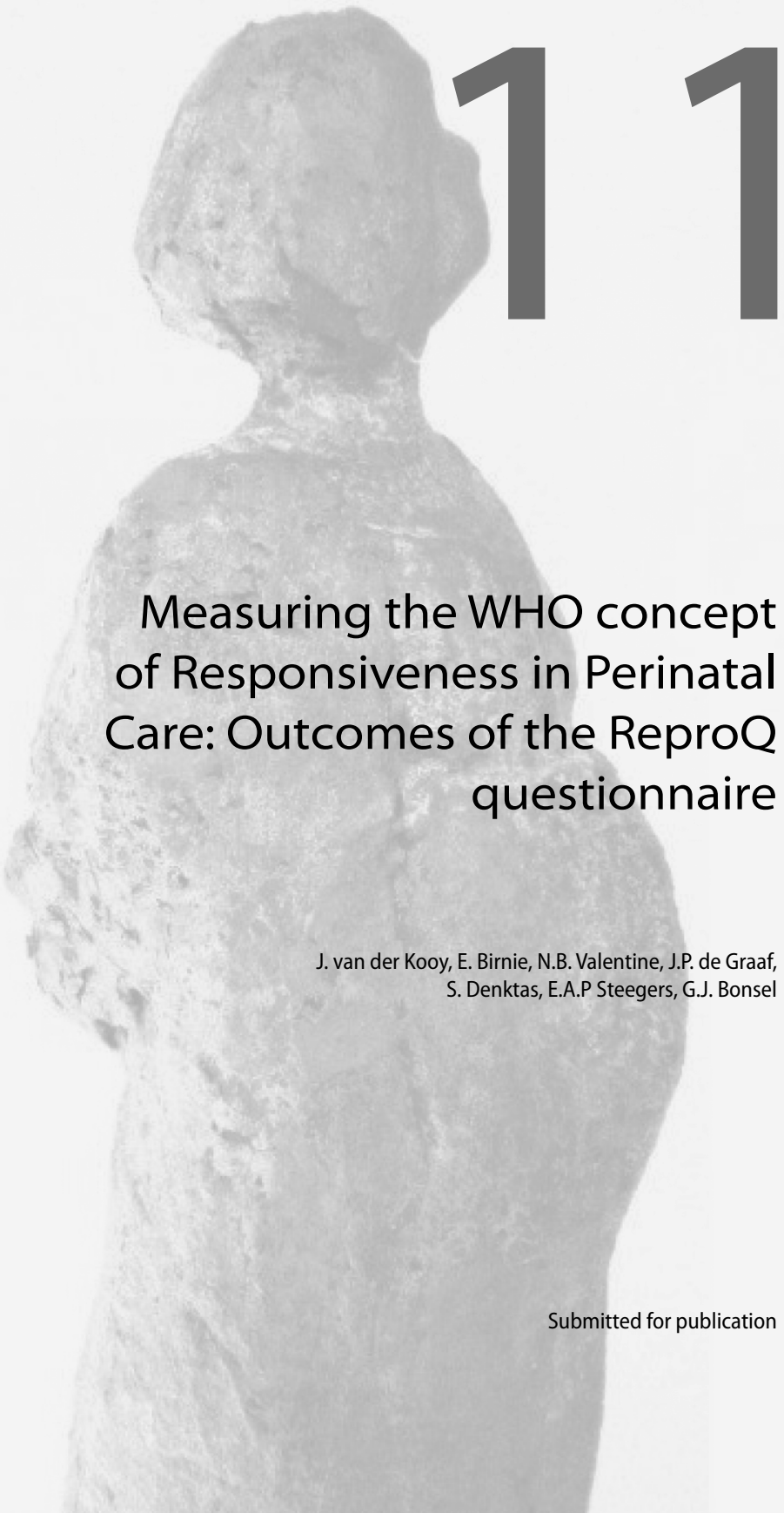
Women with a psychiatric history mentioned Confidentiality as second most important domain followed by Communication in third place. This could be due to the fact that mental health problems remain a taboo and confidentiality is a critical part of care provided to persons with mental illness.[20] Reasons for these contrary results are yet to be explored. A second remarkable finding was that women who visited the gynaecologist for antenatal care attached more value to Choice and Continuity. In the Dutch perinatal system pregnant women often change gynaecologist every visit, as antenatal care does not allow for just one person as caregiver.

Conclusion

This study shows that subgroup characteristics do not have a significant effect on the relative importance of domains, since individual heterogeneity exists. Therefore we conclude that all perinatal subpopulations can use the same set of equally valued quality ('responsiveness') domains. Herewith judgements on caregivers can be made across heterogeneous subpopulations.

REFERENCES

1. Valentine NB, de Silva A, Kawabata K, Darby C, Murray CJL, Evans DB (eds). Health system responsiveness: concepts, domains and measurement. Geneva: World Health Organisation; 2003.
2. Murray CJ, Frenk J. A framework for assessing the performance of health systems. Bull World Health Organ. 2000;78(6):717-31.
3. de Silva A (ed). A framework for measuring responsiveness. Global Programme on Evidence for Health Policy Discussion Paper. No. 32. Geneva: World Health Organisation.
4. Gostin L (ed). The domains of health responsiveness: a human rights assessment. Geneva: World Health Organisation; 2002.
5. Valentine N, Darby C, Bonsel GJ. Which aspects of non-clinical quality of care are most important? Results from WHO's general population surveys of "health systems responsiveness" in 41 countries. Soc Sci Med. 2008 May;66(9):1939-50.
6. Bramesfeld A, Wedegartner F, Elgeti H, Bisson S. How does mental health care perform in respect to service users' expectations? Evaluating inpatient and outpatient care in Germany with the WHO responsiveness concept. BMC Health Serv Res. 2007;7:99.
7. Liabsuetrakul T, Petmanee P, Sanguanchua S, Oumudee N. Health system responsiveness for delivery care in Southern Thailand. Int J Qual Health Care. 2012 Apr;24(2):169-75.
8. Rashidian A, Kavosi Z, Majdzadeh R, et al. Assessing health system responsiveness: a household survey in 17th district of tehran. Iran Red Crescent Med J. 2011 May;13(5):302-8.
9. Valentine NB, Bonsel GJ, Murray CJ. Measuring quality of health care from the user's perspective in 41 countries: psychometric properties of WHO's questions on health systems responsiveness. Qual Life Res. 2007 Sep;16(7):1107-25.
10. Schmidt R. Managing Delphi Surveys Using Nonparametic Statistical Techniques. Decision Sciences. 28: p. 763-774. 1997.
11. Sapsford RA, A. . Research Methods for Nurses and Caring professions. Buckinham: Open University Press. 1992.
12. Smith JA, Scammon DL, Beck SL. Using patient focus groups for new patient services. Jt Comm J Qual Improv. 1995 Jan;21(1):22-31.
13. Boshuizen HC, Viet AL, Picavet HS, Botterweck A, van Loon AJ. Non-response in a survey of cardiovascular risk factors in the Dutch population: determinants and resulting biases. Public Health. 2006 Apr;120(4):297-308.
14. Gayet-Ageron A, Agoritsas T, Schiesari L, Kolly V, Perneger TV. Barriers to participation in a patient satisfaction survey: who are we missing? PLoS One. 2011;6(10):e26852.
15. Salomon JA, Murray CJ. A multi-method approach to measuring health-state valuations. Health Econ. 2004 Mar;13(3):281-90.
16. Craig BM, Busschbach JJ, Salomon JA. Keep it simple: ranking health states yields values similar to cardinal measurement approaches. J Clin Epidemiol. 2009 Mar;62(3):296-305.
17. Sitzia J, Wood N. Patient satisfaction: a review of issues and concepts. Soc Sci Med. 1997 Dec;45(12):1829-43.
18. Letkovicova H. The Health Systems Responsiveness Analytical Guidelines for Surveys in the Multi-Country Survey Study. World Health Organisation: Geneva. . 2005.
19. Wensing M, Jung HP, Mainz J, Olesen F, Grol R. A systematic review of the literature on patient priorities for general practice care. Part 1: Description of the research domain. Soc Sci Med. 1998 Nov;47(10):1573-88.
20. Roberts LW, Geppert CM, Bailey R. Ethics in psychiatric practice: essential ethics skills, informed consent, the therapeutic relationship, and confidentiality. J Psychiatr Pract. 2002 Sep;8(5):290-305.
21. Sutherland HJ, Lockwood GA, Minkin S, Tritchler DL, Till JE, Llewellyn-Thomas HA. Measuring satisfaction with health care: a comparison of single with paired rating strategies. Soc Sci Med. 1989;28(1):53-8.



11

Measuring the WHO concept of Responsiveness in Perinatal Care: Outcomes of the ReproQ questionnaire

J. van der Kooy, E. Birnie, N.B. Valentine, J.P. de Graaf,
S. Denktas, E.A.P Steegers, G.J. Bonsel

Submitted for publication

ABSTRACT

Background The concept of responsiveness was introduced by the World Health Organization (WHO) to address service quality in an international comparable way. Responsiveness is defined as aspects of the way individuals are treated and the environment in which they are treated during health system interactions.

Objectives The aim of this study is to assess responsiveness outcomes during perinatal care using a newly developed questionnaire based on the WHO concept, the ReproQ.

Method The ReproQ was developed between October 2009 and February 2010 by translating the eight-domain WHO concept and adjusting it to perinatal care. Women from the Netherlands were recruited two weeks post partum. We investigated outcomes stratified for the antenatal and delivery phase. Poor outcome represents respondents who reported 'very bad', 'bad' or 'moderate' outcome on an item. When over 33% of the items were rated poor, the whole domain was judged as poor.

Results A total of 171 women consented to participation. Responsiveness poor outcome ranges from 5.9% to 31.7% within the antenatal phase and from 9.7% to 27.1% within the delivery phase. Within both phases domains covering the 'respect for persons' category are judged better than the domains covering the 'client orientation'. Parity, marital status and ethnicity influenced responsiveness outcome in the antenatal phase. Obstetric history and adverse events (e.g. intervention, hospital admission) mainly influence responsiveness in the delivery phase. The assigned importance of each domain drawn against its current overall performance showed almost a linear line, showing a better responsiveness outcome in the more important domains.

Conclusion Generally, responsiveness of the perinatal care system in the Netherlands performs quite well in absolute terms. Responsiveness of the four 'client orientations' domains underperform compared to the 'respect to persons' domains. Background characteristics show little systematic impact on responsiveness. Generally, domains which were found to be more important ('respect to persons' domains) had a better responsiveness outcome.

INTRODUCTION

The performance of perinatal care is often judged by endpoints such as perinatal morbidity and mortality and costs. However, quality of care literature supports the view that non-clinical aspects of health care, such as service quality, are important aspects of the system's performance too and, moreover, may affect clinical outcomes.[1-3] Better service quality is thought to increase compliance with medical treatment, and to improve information transfer and utilization of health services.[4-7] Governments of Western countries increasingly acknowledge the importance of incorporating service quality when the performance of the system is monitored.[8-9]

The concept of responsiveness was introduced by the World Health Organization in 2000 as one of the available approaches to address service quality in an international comparable way. In addition, the concept enables trade-off between clinical quality and service quality since it is based on utility theory.[10] Responsiveness is defined as aspects of the way individuals are treated and the environment in which they are treated during health system interactions.[10] It contains non-financial, non-clinical qualities of care that reflect respect for human dignity and interpersonal aspects of the care process. Human rights law argues that the responsiveness features of a health system are important in their own right.[10-12]

Perinatal care in the Netherlands can be regarded as a chain, which covers antenatal (outpatient) care, delivery (outpatient/inpatient care) and post partum (inpatient) care. Perinatal health care is provided by independently operating community midwives providing care for low-risk pregnant women (primary healthcare) and gynecologists providing in-hospital care for high-risk women (secondary and tertiary care). Referrals occur throughout the whole continuum of pregnancy. Most women receive post partum care by a community midwife.

The performance of the perinatal health care system in the Netherlands has come under scrutiny since the national perinatal mortality rate showed to be one of the highest in Europe.[13-15] Several initiatives to improve the perinatal health care system are made and evaluation of non clinical aspects of these health system changes one of the goals.[16] Sofar only few attempts have been made to evaluate the non-clinical aspects of quality of the perinatal health care system such that not only the pluriformity of the perinatal chain is covered, but also that international comparison is possible.[17-19] The aim of this study is to assess responsiveness of perinatal health care system in the Netherlands using the newly developed ReproQ questionnaire based on the WHO concept. We investigated the responsiveness outcomes stratified for the antenatal phase (the period from the onset of pregnancy until the onset of delivery), and the delivery phase (actual delivery).

Responsiveness outcomes on postpartum phase are not presented in this paper. We assessed the influence of background characteristics on responsiveness outcomes. And in addition, we assessed whether more important valued domains perform better.

METHODS

Questionnaire

The ReproQ questionnaire was developed to assess the responsiveness of perinatal health care system in the Netherlands. ReproQ is based on the same eight domains identified in WHO's review of the patient satisfaction and quality of care literature, i.e. Dignity, Autonomy, Confidentiality, Communication (labeled as the 'respect for persons' domains), Prompt Attention, Social Consideration, Quality of Basic Amenities, and Choice and Continuity (labeled as 'client orientation' domains).

These domains are claimed to be of universal importance in all health systems, during any client-system interaction (including personal and non-personal health services) and for the population's interaction with insurers and other administrative bodies of the health system. While it is recognized that persons may differ regarding the relative importance of each domain, and that specific domains may be of extra relevance in particular health care interactions, it is assumed that the quality of any interaction is sufficiently covered by these eight domains.

The WHO developed a survey to assess the responsiveness in an international comparable way, which was administrated between 2000-2001 as part of the Multi-Country Survey Study on Health and Health Systems Responsiveness (MCS Study) and again in 2002-2003 as part of the World Health Survey (WHS).[10,20]

ReproQ, as described here, was developed between October 2009 and February 2010, and its questions were derived from these WHO questionnaires. The ReproQ asks the same questions for the three different phases of perinatal care: antenatal phase (the period from the onset of pregnancy until the onset of delivery), delivery phase and post partum phase (covering the first ten days after childbirth). Rather than pointing to a single event, or last visit, we assumed that it is more relevant to judge all visits during the antenatal care period rather than just a single visit that may be biased in either way by a particular incident. The same applies to the postnatal maternity period. An event-like approach over delivery, however, seems natural. The professionals and settings where the items referred

to were made specific for the Dutch perinatal pare system (e.g. 'doctor' was translated into 'midwife' or 'gynecologist'). When two health care professionals are involved (e.g. 'midwife' and 'nurse'), similar questions within each domain were repeated for each health care professional separately.

Each phase was covered by the above mentioned eight domains, with 2-7 items per domain. The standardized response mode consisted of 5 options: 'very good', 'good', 'moderate', 'bad', and 'very bad'. The ReproQ consisted of 104 responsiveness questions distributed over three phases (25 antenatal, 40 delivery, 39 postpartum phase). Twenty-nine items on maternal and health care characteristics were added.

Questions from the WHO questionnaire were translated into Dutch according to a predefined protocol. First, questionnaires were translated by the research team. Expert meetings consisting of gynecologists, midwives, nurses, public health experts and researchers were held to judge the translation and comprehensiveness of the item list. Many among these professionals had working experience in English speaking countries. Next, backward translation of each question was performed and comparison was made with the original English questionnaires. Improvements were made and final consensus was reached on each question. A question was added asking which domain participants judged as most important.

The completeness of domains was judged in terms of being comprehensive (are all non-clinical areas covered, which clients and professionals put forward either as positive experience or negatively as complaint), and in terms of being balanced (have all domains included about equal importance). For each domain the candidate pool of items was checked whether each item fitted the domain definition sufficiently. As candidate items could differ per phase, this was discussed for each phase separately (e.g. the item 'quality of the food' during antenatal visits was excluded). Finally in view of the forthcoming system changes, we asked experts to check whether the domains would remain valid under ongoing and anticipated health system changes. All stakeholders agreed on the final list that the stated requirements were met.

Six primiparous and multiparous pregnant women were invited to judge the feasibility of the draft version of the questionnaire. Figure 11.1 shows the eight domains and items for the antenatal phase.

The ReproQ interview-based questionnaire demonstrated satisfactory psychometric properties, with the potential to discriminate between quality of care levels (van der Kooy et al, unpublished data).

Dignity	Were physical examinations and treatments done in a way that respected your privacy?
	Did the examination rooms ensure your privacy?
	Were you treated with respect by your health care provider?
Autonomy	How well were you involved in making decisions regarding your examinations or treatments?
	Were you able to refuse examinations or treatments?
	Were you asked permission before testing or starting treatment?
Confidentiality of Information	Were consultations carried out in a manner that protected your confidentiality?
	Was confidentiality kept on the information provided by you?
	Was your medical record kept confidential?
Communication	How well were things explained by your health care provider in a way you could understand?
	Was written information provided in such a way you could understand?
	Were you encouraged to ask questions about your health problems, treatment and care?
	Were you given time to ask questions about your health problem or treatment?
	Was information on the health service's contact, location and parking information clear to you?
Prompt Attention	How well did you receive prompt attention at your health service?
	How did you experience the waiting time after you asked for help?
	How well was the accessibility by phone?
	How do you rate the travel time to your health service?
Social Consideration	Did the health care provider facilitate the support of your relatives and friends?
	Was the home situation taken into consideration when planning an appointment?
Quality of basic amenities	How do you rate the quality of the hygiene of the toilets?
	How do you rate the overall quality of the surroundings, for example, space, seating, fresh air and cleanness?
	How do you rate the quality of the food?
Choice and Continuity of Health Care Provider	Were you able to choose your own health care provider?
	Were you able to use other health care services other than the one you usually went to?
	How well was the continuity of care by one health care provider?
	Were you able to choose your own place of delivery?

Figure 11.1 The eight domains with the items given for the antenatal phase.

Study population; data collection

The study was designed a cross sectional study. Women were recruited from three primary care midwifery practices in Rotterdam, the Netherlands, between February 2010 and March 2011. The three practices cover the north side of Rotterdam. Almost all women in

the Netherlands, regardless from whom they received antepartum and birth care, receive post partum care by a community midwife. Women or their partners were required to speak and understand Dutch sufficiently. Written informed consent was obtained.

At the post partum visit two weeks after delivery, women were invited in a consecutive order, using the day of birth, by their own midwife to participate in a 30 minute face-to-face structured interview with an independent interviewer. This interview was held at another site, usually at home. The face-to-face interviews were carried out by ten trained independent interviewers and covered questions on maternal and health characteristics and on responsiveness outcomes on the antenatal, delivery and post partum phase. Interviewees were invited to respond to all questions, yet never forced to. Study approval was granted by the Medical Ethical Committee, Erasmus Medical Centre, Rotterdam, the Netherlands, no MEC2012207.

Data handling

Records were regarded missing if scores on all phases were missing (antenatal, delivery and post partum phase). If response was partial, the response was evaluated per phase. Respondents were excluded for one phase if all items were missing for that phase. This implies that some respondents could be excluded from one phase, while being included in the other phases. Missing values within a phase were imputed with the mean when only up to 3 items were missing and selective missing among different subgroups was excluded. Variables with over 30% missing values were not imputed and excluded from analysis.

Responsiveness outcome measurements and background characteristics

For questions with a standardized response mode consisting of 5 options ('very good', 'good', 'moderate', 'bad', and 'very bad'), answers were grouped into binary categories; good and poor outcome. We choose to define responsiveness outcome poor when a respondent reported the item as 'very bad', 'bad' or 'moderate'. When over 33% of the items were rated poor within a domain, the whole domain was judged as poor. This procedure was repeated for each domain and phase separately.

Background characteristics consisted out of maternal, clinical outcome and healthcare characteristics, including parity (nulliparous/multiparous), age (<30/>30 years), ethnicity (Dutch/non-Dutch), education level (low or middle/high), marital status (single/relationship or

married), living in a deprived neighborhood (yes/no, based on 4-digit zip-codes and a public list of deprived, zip-code based, neighbourhoods issued by the Dutch government)[21], Dutch language proficiency (good/weak or poor), obstetric history (yes/no, based on self reported of mother or child outcomes which required a medical intervention by a gynaecologist), adverse child outcome (yes/no, based on self reported asphyxia, (possible) congenital anomaly, infection, small for gestational age (child too small), and/or premature birth), paediatric hospital admission (yes/no), receiving pain medication when requested (yes/no), receiving an intervention (yes/no, instrumental delivery or a caesarean section), maternal hospital admission (yes/no), day of delivery (weekend/weekday), time of delivery (8-18hr/18-8hr), healthcare pathway during pregnancy (referral to secondary care during antenatal or birth care, yes/no), perinatal healthcare pathway (Start antenatal care with midwife, not referred; Start antenatal care with midwife, referred during antenatal care to gynaecologist; Start antenatal care with midwife, referred during birth care to gynaecologist; Antenatal and birth care with gynaecologist).

Analysis

Data were analysed using SPSS software version 17.0. The unpaired Student's t-test or the Chi square test were used to compare groups on respondent's characteristics.

The impact of background characteristics on a poor responsiveness outcome was assessed for each domains using multivariate logistic regression (forward stepwise analysis; inclusion $p<0.10$; exclusion $p>0.05$). Only significant results ($p<0.05$) were presented. For each domain the percentage of respondents who valued that domain to be the most important was calculated. The assigned importance of each domain was drawn against its current overall performance (% good responsiveness).

RESULTS

A total of 274 respondents were invited for participation, 180 respondents (66%) agreed to be interviewed. Reasons for non-participation included the time burden, feeling at unease having a stranger visit their home, and logistic reasons such as incorrect phone number, or incorrect address. Of the 180 interviews planned, seven interviews (4%) were cancelled by the women and two interviews (1%) were cut short because the respondent's language proficiency was inadequate and no translator was present. The remaining 171 interviews (95%) were used for analyses. Table 11.1 shows the respondent's characteristics, pregnancy outcomes and healthcare characteristics.

Table 11.1 Respondent's characteristics, pregnancy outcomes and health care characteristics

Variable	Total 171 n	%
Maternal Age* **		
<19 years	3	2%
20-25 years	15	9%
25-34 years (REF)	119	70%
>35 years	33	19%
missing	1	1%
Parity*		
Primiparous	97	57%
Multiparous (REF)	74	43%
Education*		
Low	6	4%
Middle	75	44%
High (REF)	90	53%
Marital status*		
single	30	18%
relationship/married (REF)	141	82%
Ethnic background*		
Dutch (REF)	94	55%
Non Dutch	77	45%
Neighbourhood*		
privileged neighbourhood (REF)	84	49%
underprivileged neighbourhood	87	51%
Proficiency (speaking) Dutch*		
good/excellent (REF)	153	89%
weak/poor	18	11%
Obstetric history* ***		
Primiparous	97	57%
Multiparous, no medical history (REF)	24	14%
Multiparous, medical history	50	29%
Perinatal health care pathway*		
(1) Start antenatal care with midwife, not referred	61	36%
(2) Start antenatal care with midwife, referred during antenatal care to gynaecologist	37	22%

(*continued*)

Table 11.1 Continued

Variable	Total 171 n	%
(3) Start antenatal care with midwife, referred during birth care to gynaecologist	57	33%
(4) Antenatal and birth care with gynaecologist	16	9%
Pain medication during current pregnancy*		
No request (REF)	79	46%
No pain medication received after requesting	32	19%
Pain medication received after requesting	58	34%
Intervention in current pregnancy* *****		
No (REF)	97	57%
Yes, no emergency intervention	51	30%
Yes, emergency intervention	21	12%
Day of delivery*		
Weekend	37	22%
Weekday (REF)	134	78%
Time of delivery*		
0-8hr	45	26%
8-18hr (REF)	82	48%
18-24hr	43	25%
missing	1	1%
Adverse outcome of child* ****		
No adverse outcome (REF)	128	75%
Adverse outcome	43	25%
Hospital admission of child*		
No admission (REF)	145	85%
Admission	26	15%
Hospital admission of the mother*		
No admission (REF)	154	90%
Admission	17	10%

* p-value <0.05 (t-test).
**mean age 30 (range 18-42).
*** Obstetric history based on self reported mother or child outcomes which required intervention of a gynaecologist.
**** Ceasarean section or instrumental delivery.
***** Adverse outcome based on self reported asphyxia (shortage of oxygen), (possible) congenital anomaly, infection, small for gestational age (child too small), premature birth.

Table 11.2 Client reported poor responsiveness for each domain, for the antenatal and birth phase separately

	Antenatal Phase		Birth Phase	
Domain	N	%	N	%
Respect for Persons				
Dignity (DI)	169	5.9%	165	9.7%
Autonomy (AU)	161	18.0%	155	15.7%
Confidentiality (CF)	159	7.8%	153	11.6%
Communication (CM)	168	20.0%	166	14.2%
Client Orientation				
Prompt Attention (PA)	169	30.0%	144	20.6%
Social Consideration (SC)	164	31.7%	158	22.1%
Quality of Basic Amenities (QA)	168	22.9%	156	23.4%
Choice and Continuity (CC)	167	28.1%	162	27.1%

Table 11.2 shows the percentage of women who reported poor responsiveness at the domain level for the antenatal and the delivery phase. The proportion of poor outcome ranged from 5.9% [dignity] to 31.7% [social consideration] in the antenatal phase and from 9.7% [dignity] to 27.1% [choice and continuity] in the delivery phase. For both phases, domains covering 'respect for persons' were judged better than the domains covering 'client orientation'.

Figure 11.2 shows the reported proportion of poor outcomes for each domain by phase and by maternal, health care and outcome characteristics. In all subgroups the proportion of poor responsiveness was lower for 'respect for persons' than for client orientation. Differences between subgroups are mainly observed within the category 'client orientation'. Multiparous women tended to show poorer responsiveness outcomes on nearly all domains. The same pattern was found in women with an obstetric history (see appendix table). Ethnic differences were mainly observed within the antenatal phase where Dutch women showed poorer responsiveness outcomes. Women living in a deprived neighbourhood and those who did not speak Dutch proficiently tended to have the same responsiveness pattern (see appendix table). Furthermore, poorer responsiveness of the 'client orientation' domains were observed in women who only visited the gynaecologist in antenatal and delivery care. Poorer responsiveness of the 'respect for the patient'

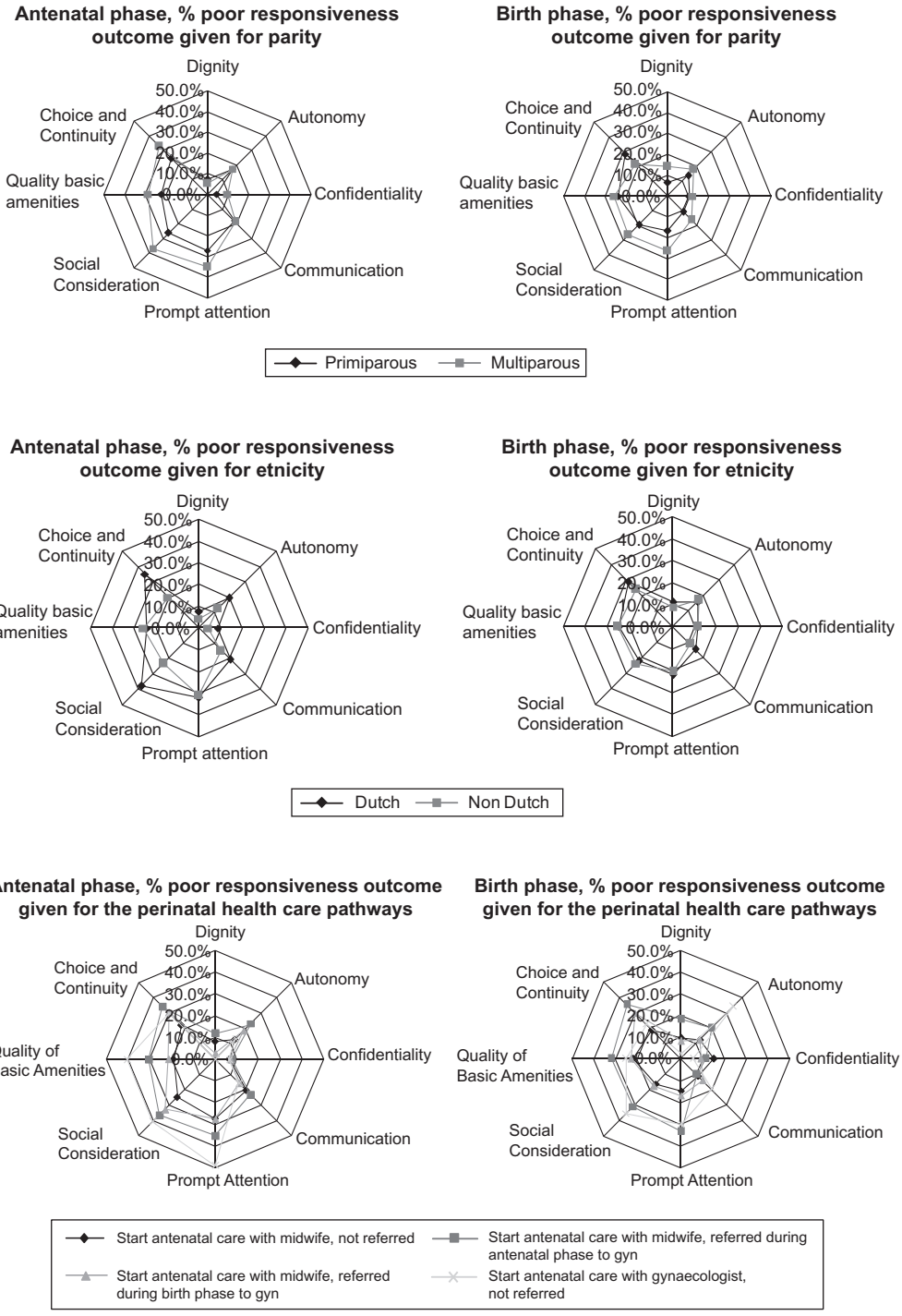


Figure 11.2 % Poor responsiveness outcomes given for maternal and health care factors.

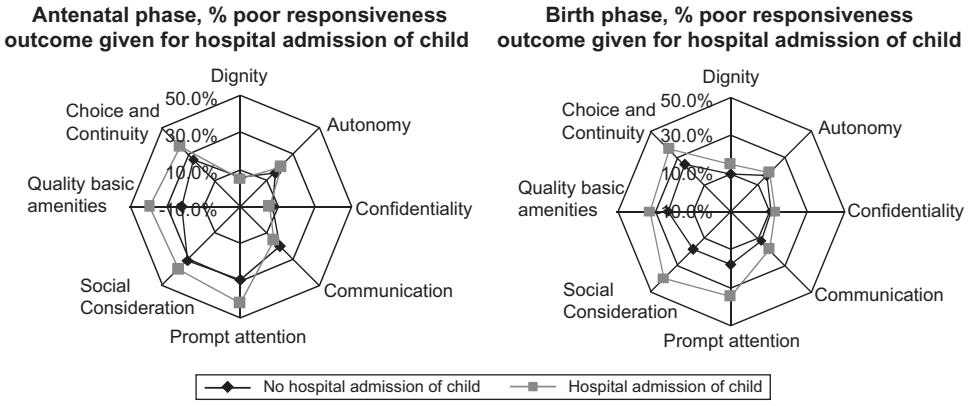


Figure 11.2 Continued

domains were observed for both the antenatal and delivery phase in women who were referred during antenatal care. Women whose child was hospitalized showed poorer responsiveness of the 'client orientation' domains. Similar patterns were observed within those who received an intervention and in those who had an adverse outcome of mother or child (see appendix table).

Table 11.3 shows the significant impact of background characteristics on poor domain responsiveness for each phase separately. Only few significant determinants were identified. The significant determinants for the antenatal phase and the delivery phase were largely different. In the antenatal phase, parity, marital status and ethnic background influenced different domains within both 'respect for persons' and 'client orientation' categories. An increased poor responsiveness in the domains covering the 'client orientation' category was found in those with an obstetric history or an adverse outcome at birth (intervention, paediatric hospital admission). This was observed in the delivery phase, and to a lesser extent in the antenatal phase.

Table 11.4 shows which domains are rated most important. The 'respect for persons' domains were generally identified as more important than the 'client orientation' domains (0.69 95%CI 0.60-0.76 vs. 0.31 95%CI 0.24-0.40).

The relationship between the proportion of good domain responsiveness and assigned domain importance was about linear, showing a better responsiveness in the more important domains except for the communication domain. (figure 11.3)

Appendix Table Percentage reporting poor outcome for each domain given for maternal, child and health care pathway outcomes

			Dignity		Autonomy		Confidentiality		Communication	
Variable	n =	%	Antenatal phase	Birth phase	Antenatal phase	Birth phase	Antenatal phase	Birth phase	Antenatal phase	Birth phase
Maternal Age* **										
<19 years	3	1.8%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	6.7%	0.0%
20-25 years	15	8.8%	2.4%	15.7%	9.5%	7.1%	7.1%	13.7%	28.6%	23.8%
25-34 years	119	69.6%	7.3%	7.3%	18.8%	16.1%	6.3%	10.8%	17.9%	12.1%
>35 years	33	19.3%	4.8%	19.3%	17.1%	25.7%	10.0%	13.0%	18.8%	11.9%
Parity*										
Primiparous	97	56.7%	7.1%	6.6%	17.1%	14.1%	4.6%	10.4%	18.4%	10.6%
Multiparous	74	43.3%	5.2%	14.3%	17.1%	18.0%	10.0%	12.0%	19.0%	16.1%
Ethnic background*										
Dutch	94	55.0%	7.7%	11.3%	19.5%	16.1%	8.5%	11.5%	20.4%	14.6%
Non Dutch	77	45.0%	4.4%	9.0%	12.5%	17.0%	4.5%	11.3%	14.6%	11.2%
Education*										
Low	6	3.5%	16.7%	23.3%	5.6%	16.7%	5.6%	0.0%	10.0%	0.0%
Middle	75	43.9%	3.6%	9.4%	16.1%	17.0%	7.8%	12.1%	18.3%	11.4%
High	90	52.6%	7.6%	9.7%	18.7%	15.7%	6.2%	11.3%	19.7%	15.3%
Marital status*										
single	141	82.5%	7.1%	10.5%	16.8%	19.6%	6.0%	12.3%	19.0%	13.8%
relationship/married	30	17.5%	2.3%	8.6%	17.8%	3.7%	10.6%	6.4%	17.5%	9.0%
Neighbourhood*										
privileged neighbourhood	84	49.1%	5.9%	6.8%	14.6%	16.3%	6.1%	7.1%	16.5%	9.2%
underprivileged neighbourhood	87	50.9%	6.4%	13.2%	19.3%	16.6%	7.6%	15.0%	20.8%	16.3%
Proficiency (speaking) Dutch*										
good/excellent	153	89.5%	6.2%	9.8%	19.6%	15.5%	7.2%	11.1%	18.5%	13.7%
weak/poor	18	10.5%	6.3%	12.5%	0.0%	23.7%	4.2%	11.8%	20.6%	6.4%
Obstetric history* ***										
Primiparous	97	56.7%	7.0%	7.0%	16.9%	15.2%	4.5%	10.5%	18.4%	10.5%
Multiparous, no medical history	24	14.0%	6.3%	7.1%	16.7%	14.5%	6.4%	6.1%	21.0%	16.1%
Multiparous, medical history	50	29.2%	4.6%	17.7%	17.4%	19.7%	11.7%	14.7%	18.1%	16.1%

Prompt attention		Social Consideration		Quality basic amenities		Choice and Continuity		TOTAL	
Antenatal phase	Birth phase	Antenatal phase	Birth phase	Antenatal phase	Birth phase	Antenatal phase	Birth phase	Antenatal phase	Birth phase
33.3%	4.8%	0.0%	0.0%	0.0%	11.1%	0.0%	7.1%	5.0%	2.9%
44.6%	26.0%	28.6%	4.8%	10.7%	22.7%	21.4%	11.8%	19.1%	15.7%
28.5%	20.5%	32.0%	22.3%	26.9%	26.6%	30.0%	11.9%	21.0%	16.0%
33.6%	22.9%	35.2%	34.9%	26.8%	22.2%	29.8%	10.9%	22.0%	20.1%
27.4%	16.7%	26.7%	19.4%	22.0%	23.8%	24.5%	28.6%	18.5%	16.3%
35.4%	26.3%	37.2%	26.4%	28.5%	25.7%	32.6%	21.7%	23.1%	20.1%
32.0%	21.9%	37.5%	22.1%	23.8%	24.5%	35.1%	28.9%	23.0%	18.9%
31.1%	20.5%	23.0%	23.8%	25.4%	25.0%	19.9%	23.9%	16.9%	17.7%
29.2%	35.7%	8.3%	16.7%	33.3%	22.2%	5.6%	11.9%	14.3%	15.8%
29.3%	18.7%	29.0%	20.2%	22.7%	24.7%	24.2%	24.0%	18.9%	17.2%
33.3%	22.1%	35.2%	25.0%	25.6%	25.2%	33.7%	29.8%	22.5%	19.3%
30.7%	21.1%	32.8%	23.1%	25.6%	26.1%	29.9%	27.0%	21.0%	19.2%
33.9%	21.4%	25.9%	18.3%	20.7%	19.7%	22.0%	23.1%	18.8%	13.8%
30.3%	15.3%	31.3%	16.8%	29.5%	23.1%	26.9%	23.5%	20.1%	14.7%
32.3%	26.5%	31.5%	27.1%	20.1%	26.6%	29.7%	28.9%	21.0%	21.3%
32.0%	21.5%	32.2%	20.9%	24.3%	25.1%	28.6%	25.9%	21.1%	18.0%
25.3%	18.0%	25.0%	31.8%	28.1%	22.9%	26.2%	29.2%	17.0%	19.5%
28.3%	17.5%	26.9%	19.1%	21.8%	24.3%	25.0%	21.8%	18.6%	15.7%
28.6%	19.9%	40.5%	18.6%	33.3%	14.4%	34.9%	28.6%	23.5%	15.7%
38.7%	29.6%	35.6%	30.1%	26.1%	31.2%	31.5%	34.9%	22.9%	24.2%

(*continued*)

Appendix Table Continued

Variable	n =	%	Dignity		Autonomy		Confidentiality		Communication	
			Antenatal phase	Birth phase	Antenatal phase	Birth phase	Antenatal phase	Birth phase	Antenatal phase	Birth phase
Adverse outcome* ****										
No adverse outcome	128	74.9%	5.4%	10.7%	16.4%	17.6%	7.1%	12.3%	17.8%	13.2%
Adverse outcome	43	25.1%	8.5%	8.4%	18.7%	12.8%	6.2%	7.9%	21.3%	12.0%
Hospital admission of child*										
No hospital admission of child	145	84.8%	6.3%	9.3%	16.2%	16.0%	7.1%	10.8%	19.4%	11.8%
Hospital admission of child	26	15.2%	5.6%	14.2%	20.8%	18.3%	5.6%	12.8%	15.0%	18.4%
Receiving painmedication when requested*										
No request	79	46.2%	7.0%	9.4%	14.1%	14.8%	5.4%	11.8%	15.3%	13.6%
No pain medication received after requesting	32	18.7%	8.1%	15.7%	23.4%	11.9%	11.6%	11.9%	21.9%	12.3%
Pain medication received after requesting	58	33.9%	3.5%	7.4%	18.5%	19.7%	6.7%	10.0%	22.2%	12.6%
Intervention* *****										
No	97	56.7%	6.1%	7.2%	26.4%	12.4%	7.6%	10.6%	18.9%	10.8%
Yes, no emergency intervention	51	29.8%	5.7%	12.5%	27.5%	19.4%	4.2%	8.3%	20.8%	14.2%
Yes, emergency intervention	21	12.3%	7.0%	17.0%	31.2%	22.9%	10.5%	20.2%	14.7%	20.4%
Hospital admission of the mother*										
No admission	154	90.1%	6.1%	10.1%	17.2%	16.1%	6.0%	11.3%	19.5%	13.9%
Admission	17	9.9%	7.7%	10.8%	14.9%	19.4%	16.7%	9.7%	9.6%	2.7%
Day of delivery*										
Weekend	37	21.6%	3.0%	5.7%	16.4%	10.7%	2.1%	7.3%	15.0%	9.4%
Weekday	134	78.4%	7.1%	11.4%	17.2%	17.8%	8.2%	12.3%	19.7%	13.8%

Prompt attention		Social Consideration		Quality basic amenities		Choice and Continuity		TOTAL	
Antenatal phase	Birth phase	Antenatal phase	Birth phase	Antenatal phase	Birth phase	Antenatal phase	Birth phase	Antenatal phase	Birth phase
30.2%	17.7%	29.7%	20.5%	22.1%	23.2%	27.6%	24.3%	19.5%	17.5%
34.4%	31.4%	36.4%	26.8%	32.1%	29.7%	30.2%	32.1%	23.5%	20.1%
29.3%	18.4%	30.2%	18.4%	21.8%	23.4%	26.7%	24.3%	19.6%	16.6%
41.7%	35.3%	37.5%	39.9%	39.6%	32.5%	36.1%	36.2%	25.2%	26.0%
28.0%	20.3%	31.0%	21.6%	21.8%	21.4%	26.4%	26.0%	18.6%	17.4%
30.4%	21.0%	29.3%	27.9%	27.6%	27.2%	27.5%	25.8%	22.5%	19.2%
36.0%	21.4%	33.3%	20.1%	28.1%	28.4%	31.2%	27.6%	22.4%	18.4%
27.2%	16.6%	33.7%	19.9%	20.5%	23.0%	22.5%	22.4%	20.3%	15.4%
38.2%	26.8%	23.8%	24.6%	34.1%	29.9%	34.3%	28.2%	23.6%	20.5%
34.2%	27.1%	36.8%	28.3%	26.3%	21.7%	40.4%	41.3%	25.1%	24.9%
32.0%	21.8%	31.5%	22.9%	24.1%	24.3%	28.4%	27.4%	20.6%	18.5%
23.8%	14.5%	30.8%	14.4%	30.8%	32.2%	26.4%	13.7%	20.1%	14.7%
34.1%	15.2%	40.0%	28.4%	18.8%	19.4%	28.9%	26.1%	19.8%	15.3%
30.5%	22.8%	29.0%	20.4%	26.3%	26.4%	28.1%	26.4%	20.8%	18.9%

(*continued*)

Appendix Table Continued

Variable	n =	%	Dignity		Autonomy		Confidentiality		Communication	
			Antenatal phase	Birth phase	Antenatal phase	Birth phase	Antenatal phase	Birth phase	Antenatal phase	Birth phase
Time of delivery*										
0-8hr	45	26.3%	3.3%	10.3%	17.6%	12.9%	5.2%	13.2%	20.1%	15.3%
8-18hr	82	48.0%	6.5%	9.4%	18.9%	19.4%	8.7%	14.0%	21.2%	14.9%
18-24hr	43	25.1%	8.9%	9.5%	13.3%	14.8%	5.5%	3.0%	12.9%	6.6%
Health care pathway*										
Start antenatal care with midwife, not referred	61	35.7%	8.2%	9.5%	12.8%	12.3%	6.5%	15.0%	20.2%	11.1%
Start antenatal care with midwife, referred during antenatal care to gynaecologist	37	21.6%	11.8%	18.1%	22.9%	19.8%	7.7%	11.3%	22.9%	10.3%
Start antenatal care with midwife, referred during birth care to gynaecologist	57	33.3%	2.6%	8.2%	18.5%	12.2%	6.8%	9.0%	15.6%	14.0%
Antenatal and birth care with gynaecologist	16	9.4%	0.0%	2.5%	13.2%	34.0%	6.5%	5.4%	15.3%	19.8%
Total	171	100.0%	6.2%	10.1%	17.0%	16.4%	6.9%	11.2%	19.9%	12.9%

* p-value <0.05 (t-test).
**mean age 30 (range 18-42).
*** Obstetric history is defined as.
**** Self reported adverse outcome defined as; asphyxie, (possible) congenital anomalie, infection, small for gestational age, large for gestational age, premature birth.
***** Ceasarean section or instrumental delivery.

Prompt attention		Social Consideration		Quality basic amenities		Choice and Continuity		TOTAL	
Antenatal phase	Birth phase	Antenatal phase	Birth phase	Antenatal phase	Birth phase	Antenatal phase	Birth phase	Antenatal phase	Birth phase
31.0%	20.9%	35.1%	22.2%	21.3%	20.1%	32.5%	29.2%	20.8%	18.0%
30.9%	18.6%	31.6%	26.1%	27.5%	27.4%	29.8%	24.2%	21.9%	19.2%
32.6%	27.1%	26.6%	15.2%	23.7%	23.8%	20.3%	27.0%	18.0%	15.9%
27.3%	14.7%	24.7%	16.2%	19.2%	21.9%	22.2%	18.8%	17.6%	14.9%
35.2%	33.4%	36.2%	31.0%	30.6%	32.1%	34.1%	35.3%	25.2%	23.9%
27.6%	17.0%	32.7%	17.8%	21.6%	23.2%	30.4%	29.8%	19.5%	16.4%
49.1%	30.2%	40.6%	35.4%	40.6%	25.1%	30.1%	22.3%	24.4%	21.8%
31.3%	21.2%	31.4%	22.2%	24.7%	24.9%	28.3%	26.3%	20.7%	18.1%

Table 11.3 The influence of background characteristics on poor responsiveness for each domain separately. (Only statistically significant results are presented)

	Antenatal Phase				Birth Phase			
Domain	Determinants	OR	95%CI		Determinants	OR	95%CI	
Dignity	Marital status	0.18	0.02	1.40	Obstetric history	2.46	0.97	6.27
	Day of delivery	0.28	0.06	1.28				
Autonomy	Ethnic background	0.52	0.26	1.06				
	Intervention	2.24	1.12	4.48				
Confidentiality	Marital status	2.64	1.00	6.92	Health care pathway	0.45	0.19	1.11
Communication	Day of delivery	0.34	0.11	1.03				
Prompt Attention	Ethnic background	0.50	0.24	1.07	Neighboorhood	2.80	1.14	6.87
	Obstetric history	2.74	1.25	6.02	Obstetric history	2.82	1.11	7.16
	Hospital Admission of Child	2.76	1.11	6.87	Adverse outcome child	2.95	1.17	7.44
	Receiving painmedication	2.54	1.20	5.40	Intervention	3.16	1.28	7.78
Social Consideration	Parity	0.42	0.21	0.83	Obstetric history	2.34	1.06	5.15
	Ethnic background	0.32	0.16	0.64	Hospital Admission of Child	3.25	1.25	8.45
Quality of Basic Amenities	Hospital Admission of Child	2.09	0.90	4.86				
Choice and Continuity	Parity	0.57	0.29	1.11	Maternal age	3.68	1.30	10.45
	Education	0.38	0.20	0.74	Parity	0.27	0.10	0.71
					Intervention	2.66	0.99	7.13
					Hospital Admission of Child	0.08	0.01	0.80
					Time of delivery	4.29	1.53	12.07

Inclusion $p<0.10$; exclusion $p<0.05$.
Determinants included (reference):
Maternal age (<30year), Parity (multiparous), Ethnic background(Dutch), Education(high), Marital status(relationship/married), Neigbourhood(priveleged), Proficiency (speaking) Dutch(good/excellent), Obstetric history(No), Adverse Outcome(No), Hospital Admission of Child(No), Receiving painmedication(No), Intervention(No), Hospital Admission of the Mother(No), Day of delivery(weekday), Time of delivery(8-18hr), Health care pathway(Not referred).

Table 11.4 Percentage reporting the domain to be most important

Domain	%
Respect For Persons	
Dignity (DI)	21.6%
Autonomy (AU)	9.6%
Confidentiality (CF)	11.2%
Communication (CM)	26.4%
Client Orientation	
Prompt Attention (PA)	10.4%
Social Consideration (SC)	4.0%
Quality of Basic Amenities (QA)	10.4%
Choice and Continuity (CC)	6.4%

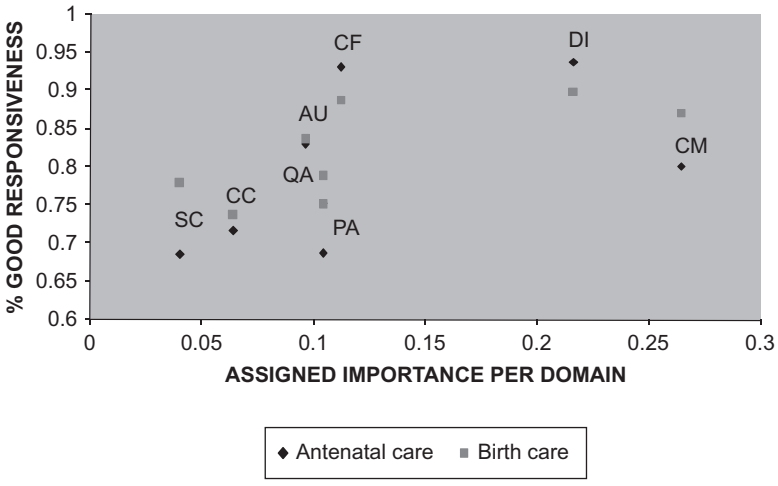


Figure 11.3 Assigned importance per domain drawn against % good responsiveness.

DISCUSSION

Generally, responsiveness of the perinatal care system in the Netherlands performs quite well in absolute terms. Responsiveness of the four 'client orientations' domains underperform compared to the 'respect to persons' domains. Background characteristics show little systematic impact on responsiveness. Parity, marital status and ethnicity influence responsiveness in the antenatal phase. Obstetric history and adverse events

(e.g. intervention, hospital admission) mainly influence responsiveness, as expected, in the delivery phase. Generally domains which were found to be more important ('respect to persons' domains) had a better responsiveness outcome.

Strengths and limitations

Some strengths are noteworthy to mention. Firstly, 66% of the invited women agreed to participate in this study. This is an effective study sample, since a response rate of 30% has been proposed as reasonable for patient satisfaction surveys and a response rate of 50% is considered to be quite high.[13,14]

Secondly, our study covered many subpopulations in Rotterdam, also subpopulations which are often missed in satisfaction surveys. More frequent among non-participants in satisfaction studies are having a language barrier, a psychiatric history, a low social economic status, a low educational level, no paid work and Muslim people.[15,16] Since our study covered these subpopulations, its generalizability is more presumably for women in perinatal care. Thirdly, interviews were conducted in such a way that known factors influencing respondent's health responsiveness outcomes were diminished as much as possible. Interviews were performed by independent interviewers, respondents were interviewed at their own homes and interviewed beyond two weeks post partum. Previous studies have shown that women who answer surveys at home are more critical compared with respondents who are interviewed in the hospital, since the latter are loyal to the institution.[22] Women being interviewed within two weeks also tend to be less critical.[23]

A few limitations merit discussion. Firstly, since only people from urban areas participated in this study, the study population is presumably representative for Dutch urban areas, but the generalizability to the whole Dutch population remains uncertain. Secondly, translation could only be arranged for some of the women who did not understand the Dutch language sufficiently, this was done by a family member of the women. This could introduce a translation bias since this was not done by a professional translator. Thirdly, all non-Dutch ethnic groups were grouped resulting in a heterogeneous subpopulation. Responsiveness outcomes in these subpopulation may differ, since studies showed that ethnicity can be of influence.[20] Fourthly, no analysis was performed on non-participants. Fifthly, carry over effects on health responsiveness outcomes within the antenatal phase cannot be excluded, since birth outcome determinants significantly influenced outcomes within the antenatal phase.

We observed poorer outcomes on the category 'client orientation', similar outcomes were found by Liabsuetrakul et al [24] This might be, because domains covering this category are in general less easier to influence than domains covering the 'respect to person' category.

The latest category could be influenced by only one professional instead of changes in the organization of care. A second explanation might be that the domains in this category are judged as more important by the health professionals and thus more attention is given.

Parity and ethnic characteristics influenced responsiveness outcome in the antenatal phase only (except for the domain choice and continuity in the delivery phase). Obstetric history and an adverse event (receiving an intervention, hospital admission of mother/ child) influenced responsiveness outcome in the delivery phase. This is to a degree in line with what was found in the CAPHS patient experience survey. They observed age, general health, education, individual health plan, and less influential ethnicity, gender and time in insurance plan to influence responses on patient experience.[25] Although we did not assess the impact of health plan and time in insurance plan, we observed similar characteristics for both phases to be of influence on responsiveness outcome, since age and parity compete with each other. Other studies that assessed patient characteristics on (some) of the WHO responsiveness domains showed similar patterns for parity, ethnicity, education and marital status. However these studies did not include birth outcomes within their analysis.[24,26]

Being referred during pregnancy seems to have a negative influence on responsiveness compared to staying with the midwife only. However, after adjusting for background characteristics we only observe referral to significantly influence the responsiveness of domain 'confidentiality' within the delivery phase. This is in line with other studies who found no association with being referred and responsiveness domains.[17,19]

The domains communication and dignity were most frequently identified as most important. This is partly in contrast with the population based survey conducted by the WHO [27] and results by Liabsuetrakul et al, who assessed the importance of responsiveness domains in Thailand.[24] They both found prompt attention and dignity to be the most important domains, followed by communication in third place. The preference for Prompt Attention may be due to the fact that in the WHO surveys Prompt Attention was operationalized in terms of geographical access and access in case of emergencies. In our study Prompt Attention focussed upon waiting times. Results from other studies which also focussed upon waiting times support our results that Prompt Attention was than valued as less important.[2-3] Bramesfeld et al saw a similar ranking, but observed a difference in ranking between in- and outpatient mental care. Hereby, observing prompt attention to be more important in outpatient care.[28]

We observed a relationship between the proportion of good domain responsiveness and assigned domain importance was about linear, showing a better responsiveness in the more important domains. Professionals might also judge these domains as most important, and therefore add more value to them.

Overall, our ReproQ questionnaire, which was derived from the WHO concept of responsiveness, demonstrated satisfactory responsiveness outcomes of the perinatal care system in the Netherlands. Based on results of our study we recommend that when evaluating the responsiveness outcomes of the perinatal health care system, antenatal care should be evaluated before the start of delivery to prevent the prevent carry over effects of birth outcomes. To improve responsiveness outcomes of the Dutch Perinatal Care system, caregivers should focus on domains covering the category ´client orientation´.

REFERENCES

1. Campbell SM, Roland MO, Buetow SA. Defining quality of care. Soc Sci Med. 2000 Dec;51(11): 1611-25.
2. Sitzia J, Wood N. Patient satisfaction: a review of issues and concepts. Soc Sci Med. 1997 Dec;45(12):1829-43.
3. Wensing M, Jung HP, Mainz J, Olesen F, Grol R. A systematic review of the literature on patient priorities for general practice care. Part 1: Description of the research domain. Soc Sci Med. 1998 Nov;47(10):1573-88.
4. McPake B. User charges for health services in developing countries: a review of the economic literature. Soc Sci Med. 1993 Jun;36(11):1397-405.
5. Ware JE, snyder MK, Wright WR, Davies AR (eds). Defining and measuring patient satisfaction with medical care. Evaluation and Program Planning. Volume 6: 247-263; 1983.
6. Wouters A. Essential national health research in developing countries: heath care financing and the quality of care. Internation Journal of Health Planning and Management. 1991;6: 253-271.
7. Gilson L, Alilio M, Heggenhougen K. Community satisfaction with primary health care services: an evaluation undertaken in the Morogoro region of Tanzania. Soc Sci Med. 1994 Sep;39(6): 767-80.
8. Greener I. Towards a history of choice in UK health policy. Sociol Health Illn. 2009 Apr;31(3):309-24.
9. Wensing M, Elwyn G. Methods for incorporating patients' views in health care. BMJ. 2003 Apr 19;326(7394):877-9.
10. Valentine NB, de Silva A, Kawabata K, Darby C, Murray CJL, Evans DB (eds). Health system responsiveness: concepts, domains and measurement. Geneva: World Health Organisation; 2003.
11. de Silva A (ed). A framework for measuring responsiveness. Global Programme on Evidence for Health Policy Discussion Paper. No. 32. Geneva: World Health Organisation.
12. Gostin L (ed). The domains of health responsiveness: a human rights assessment. Geneva: World Health Organisation; 2002.
13. Mohangoo AD, Buitendijk SE, Hukkelhoven CW, et al. [Higher perinatal mortality in The Netherlands than in other European countries: the Peristat-II study]. Ned Tijdschr Geneeskd. 2008 Dec 13;152(50):2718-27.
14. Ravelli AC, Tromp M, van Huis M, et al. Decreasing perinatal mortality in The Netherlands, 2000-2006: a record linkage study. J Epidemiol Community Health. 2009 Sep;63(9):761-5.
15. Zeitlin J, Wildman K, Breart G, et al. PERISTAT: indicators for monitoring and evaluating perinatal health in Europe. Eur J Public Health. 2003 Sep;13(3 Suppl):29-37.
16. van der Velden J. Een goed begin. Veilige zorg rond zwangerschap en geboorte. http://www.knov.nl/docs/uploads/Advies_stg_zw_en_geb_Een_goed_beginpdf_spread-versie_voor_media.pdf, 2009.
17. Christiaens W, Gouwy A, Bracke P. Does a referral from home to hospital affect satisfaction with childbirth? A cross-national comparison. BMC Health Serv Res. 2007;7:109.

18. Wiegers TA. The quality of maternity care services as experienced by women in the Netherlands. BMC Pregnancy Childbirth. 2009;9:18.

19. Wiegers TA, van der Zee J, Keirse MJ. Transfer from home to hospital: what is its effect on the experience of childbirth? Birth. 1998 Mar;25(1): 19-24.

20. Valentine NB, Bonsel GJ, Murray CJ. Measuring quality of health care from the user's perspective in 41 countries: psychometric properties of WHO's questions on health systems responsiveness. Qual Life Res. 2007 Sep;16(7):1107-25.

21. The MolatEVGdodnatpScftdodn, Hague: VROM IDAoawrnd-e-pkb-k-w-o-r-v-a-e-p-p.

22. Lumley J. Assessing satisfaction with childbirth. Birth. 1985 Fall;12(3):141-5.

23. van Teijlingen ER, Hundley V, Rennie AM, Graham W, Fitzmaurice A. Maternity satisfaction studies and their limitations: "What is, must still be best". Birth. 2003 Jun;30(2):75-82.

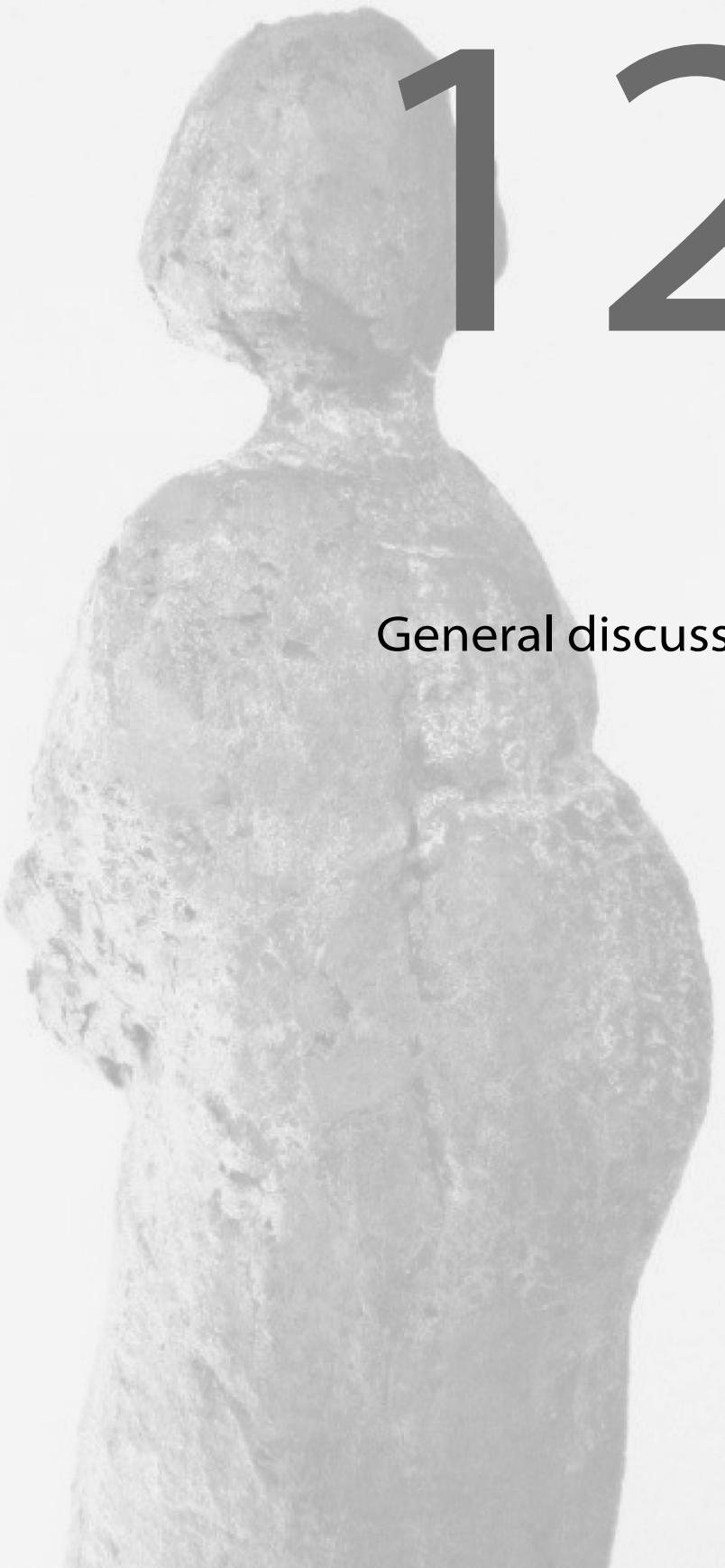
24. Liabsuetrakul T, Petmanee P, Sanguanchua S, Oumudee N. Health system responsiveness for delivery care in Southern Thailand. Int J Qual Health Care. 2012 Apr;24(2):169-75.

25. Zweifler J, Hughes S, Lopez RA. Controlling for race/ethnicity: a comparison of California commercial health plans CAHPS scores to NCBD benchmarks. Int J Equity Health. 2010;9:4.

26. Raleigh VS, Hussey D, Seccombe I, Hallt K. Ethnic and social inequalities in women's experience of maternity care in England: results of a national survey. J R Soc Med. 2010 May;103(5):188-98.

27. Valentine N, Darby C, Bonsel GJ. Which aspects of non-clinical quality of care are most important? Results from WHO's general population surveys of "health systems responsiveness" in 41 countries. Soc Sci Med. 2008 May;66(9):1939-50.

28. Bramesfeld A, Wedegartner F, Elgeti H, Bisson S. How does mental health care perform in respect to service users' expectations? Evaluating inpatient and outpatient care in Germany with the WHO responsiveness concept. BMC Health Serv Res. 2007;7:99.



12

General discussion

AIMS OF THIS THESIS

The merits of the Dutch perinatal care system have come under scrutiny since the national perinatal mortality rate showed to be one of the highest in Europe.[1-3] The functioning of this unique system depends on close cooperation of the health care professionals, availability of (different) facilities, no financial barriers, and well informed users.

This thesis describes an assessment of the Dutch perinatal care system focusing on both clinical and non clinical aspects, and evaluates possible innovative strategies to improve the Dutch perinatal care. It also introduces some new tools for evaluation in this context. Throughout this thesis, the assessment of the Dutch perinatal care system focuses on risk selection and midwife-led birth care, using an observational design and data from 2000-2007.

The main objectives of this thesis were;

I. What are possible explanations for the underperformance of the Dutch perinatal care system, in particular focussing on risk selection and midwife-led birth care?

II. How can Dutch perinatal care be improved, in particular, are there innovative strategies available to improve risk selection and midwife-led birth care?

III. How can the performance of the Dutch perinatal care be evaluated on non clinical aspects of quality of care?

This chapter discusses methodological considerations, presents the findings of this thesis in a broader perspective, and offers recommendations.

Throughout this chapter we used the term Big4 and Big3 as an indication for manifest high risk pregnancies. Big4 pregnancies are defined as: congenital abnormalities (list defined), intrauterine growth restriction (SGA, birth weight below the 10th percentile for gestational age, gender and parity specific), preterm birth (< 37th week of gestation) or low Apgar score (< 7, measured 5 minutes after birth). Big3 pregnancies are defined as: congenital abnormalities (list defined), intrauterine growth restriction (SGA, birth weight below the 10th percentile for gestational age, gender and parity specific), or preterm birth (< 37th week of gestation). We assume these conditions can be measured with sufficient reliability.

MAIN FINDINGS

Part I. What are possible explanations for the underperformance of the Dutch perinatal care system, in particular focussing on risk selection and midwife-led birth care?

We observed that the Dutch (two-tier) perinatal care system in 2000-2007 data insufficiently separates low from high risk pregnancies, and that a considerable part of this selection is happening during parturition rather than well before. This insufficiency in our view in part reflects the poor midwife-obstetrician interaction and the inadequacy of putting existing knowledge and tools to each case. Our studies offer evidence that organisational aspects rather than disadvantageous maternal characteristics alone, e.g. older maternal age, smoking and multiple pregnancies, are a major driver of the poor Dutch results.[1,4-6] Underlying patterns and long term effects of insufficient risk selection on neonatal and maternal outcome are yet to be explored.

Improving current risk selection, especially preconceptionally or antenatally, would lead to early detection and possible prevention of medium and high risk pregnancies, and may lead to an improvement of Dutch perinatal outcomes. One possible strategy to improve current risk selection is the increase of midwives' competence and capabilities. Another one is the introduction of up to date screen methods (like R4U) at onset of pregnancy, complemented by active systematic prevention (including non-medical risks) and more advanced monitoring (e.g. ultra sound) at later pregnancy stages. The success of these strategies depends on a joint collaboration of midwives and obstetricians. Indeed, we believe that the required pace of change depends most on the adoption of 'shared care'; better cooperation between midwives and obstetricians who are jointly responsible for the determination and monitoring of a woman's risk status.[7] Shared obstetric care has already been implemented in some form in other Western countries such as Australia and the United Kingdom.[8-10] One study demonstrated a 27% increase in the detection rate of intrauterine growth restriction for women receiving shared obstetric care as opposed to conventional obstetric care.[11]

Although planned home births showed the advantages of lower intervention rates, especially for multiparous women, this advantage is counterbalanced by the possible additional mortality in undetected or unexpected Big3 pregnancies. This pattern of

outcomes may affect choices and the outcome on an individual level, but setting effects per se do not give an explanation for the poor performance of the Dutch perinatal care system in general, since absolute rates and observed differences are rather small between planned home and hospital births.

We are aware that self selection of pregnant women planning their home either at home or in the hospital can coincide with implicit or explicit selection by the midwife who may tend to 'refer' to hospital if she feels uncomfortable with the risk level at home. The choice for interventions in planned home and hospital births partly suggests better fit of intervention patterns to risk profile, however deeper insight in how these patterns arise are yet to be explored.

Setting differences on maternal and perinatal outcome were observed analysing the Dutch perinatal care system in general. However, quality of care may differ among different settings and practices. This was not analysed since practice specific analysis is not allowed.

Part II. What are possible strategies to improve the Dutch unfavourable position?

The introduction of a midwife-led birth centre, with the described philosophy and organization, affects existing working procedures and organizational structures, such as the collaboration of different disciplines, the joint development of protocols, etc. It affects adverse perinatal and maternal outcomes directly, but most likely in indirect ways too.

We observed that the introduction of a midwife-led birth centre redistributed women according to place of midwife-led delivery. Surprisingly, woman planning their birth in the midwife-led birth centre had the most unfavourable risk profile compared to women planning their midwife-led birth in the hospital or at home. A similar trend was observed in the Birthplace cohort[12], but most other international studies showed the opposite.[13-16]

Differential use of these options can be explained by several factors, either intentional or coincidental. After the introduction of the birth centre, low risk women could not plan their delivery in the hospital adjacent the birth centre anymore, but are still able to plan their delivery in other nearby hospitals. This may have led to a shift from the previously planned hospital births to the birth centre. Secondly, our birth centre aims to provide risk led care, with special attention to ethnic minorities and women with a low social economic background. This encourages caregivers to offer the higher risk women (among the low risk group) more explicitly the option of a birth centre delivery. Furthermore, in contrast to planned hospital births, women can receive postpartum care for at least four days in the midwife-led birth centre as an option. This may also attract both higher risk women and their caregivers.

After the introduction of the birth centre, intrapartum and early neonatal mortality and morbidities tended to decrease, while overall intervention rates were unaffected. Its introduction seems to benefit the outcome of midwife-led deliveries, which suggest improved better care through more adequate risk selection even at this stage. However, its introduction will not have a large effect on perinatal mortality at large, since mortality rates are low and only small differences are observed, and since about 66% of perinatal mortality is stillbirth.

Overall intervention rates were not affected by the introduction of the birth centre. While the underlying pattern suggests a better choice for interventions relating to risk profiles, the introduction of the midwife-led birth centre did not represent an up- or downward pressure towards intervention rates in general. Previous studies on birth centres showed lower intervention rates combined with an equal, or even better, performance.[14-15,17] These studies, however, did not or only partially adjust for case mix differences.[14-15,17] The few available randomized controlled comparisons also showed lower rates[18-19], or at most an equal intervention rate with equal perinatal outcomes.[13,20-21] As these trials suffered from non participation or small study size [13,18-21], and showed difficulty to combine results [22], our study provides observational evidence in a large unselected cohort that overall intervention rates are not affected by the introduction of a midwife-led birth centre.

We believe that changes in existing working procedures and the general organization, as described before, are responsible for positive change in perinatal and maternal outcomes. Key concepts are integral organization, risk led care, dedicated care during transitions (in case of admission, referral) and protocollized risk communication and application of risk adjusted (standardized) protocols.

It was beyond the scope of this thesis and the size of the data collection to relate all specific organisational changes to specific effects. The birth centre, however, allowed for studying the use of R4U as an adjuvant triage tool to better detect SGA at the start of birth in an assumed low risk population, and showed promising results. Using the R4U preconceptionally and antenatally, particularly in combination with advanced screening tests (e.g. ultra sound) may lead to even better detection of medium and high risk pregnancies at an early stage. Early detection gives opportunities for intervention and leads to an adequate place of birth. Perinatal and maternal outcomes will likely benefit from this on a large scale. This is currently evaluated in a national program named Health Pregnancy for All.[23]

The findings in our study also provide more insight in underlying patterns leading to SGA. Better detection of SGA at birth may lead to better individual outcomes, since if anticipated SGA pregnancies receive specialized care with e.g. alerted neonatologists present.

The third innovation studied in this thesis addresses the reintroduction of N_2O analgesia. Occupational safety, patient safety and the effect on pain are key features of the evaluation of N_2O analgesia in this context. Occupational safety could be attained; this relied on the strict and complete application of preventive strategies. Reported rates of complications of N_2O administration were low in our study. Rosen et al. concluded in his review that administration of N_2O during labour does not affect the course of labour, mother or child outcomes.[24] However, long term effects on child outcomes should be explored.

In our study the effect on pain was could be challenged since VAS scores did not decrease to the extent reported if women receive epidural analgesia. But here it can be argued that a pain-free delivery is not the implicit goal; the effect on pain can be judged in terms of relief and additional measures such as the decrease of feeling nervous or anxious, patient satisfaction, and, the need for other additional analgesia. Our data on feeling nervous or anxious, the need for additional analgesia and patient satisfaction showed promising results. Additionally, evaluation of the use of N_2O analgesia includes competitive features like (1) immediate availability, (2) intermittent availability, (3) the allowance for self control by the women, (4) non invasiveness, and (5) low running costs.[25-26] Adopting the comprehensive evaluation format we strongly recommend the reintroduction of N_2O analgesia, respecting the conditions for safe provision limiting the use to birth centres adjacent to a hospital and delivery rooms in a hospital with instructed and trained personnel.

Part III. Evaluation of the current Dutch Perinatal System focussing on non clinical aspects of quality of care; responsiveness outcomes.

The ReproQ as developed in our studies showed to be a promising tool for the evaluation of service quality of perinatal care in the Netherlands, and possibly other countries. It can be used to evaluate and/or compare health care systems, settings and practices on both national and local levels since it was specifically developed to overcome system differences. It captures the patient's actual experiences and ultimately enables a quantitative trade-off between service level and clinical outcome if these vary across systems.

Responsiveness outcomes of Dutch perinatal care showed that improvements can especially be made in the 'client orientation' domains.

After we introduced the ReproQ as described, the instrument was further changed into an abridged stand alone digital questionnaire to accommodate the requirements of permanent national use in perinatal quality measurement. These requirements are endorsed by Miletus on behalf of all health insurance companies, to arrive at the national CQI instruments (see.http://www.centrumklantervaringzorg.nl/wat-is-de-cq-index.html; in Dutch).

METHOLOGICAL ISSUES

Some overall methodological issues can be discussed on the studies covered in this thesis. Since part I and part II share common methodological issues on data sources and study design, these will be discussed together.

Methological issues PART I and II

Data sources

Existing medical registries were used to answer the study questions in part I and II. Medical registries try to capture particular actions in health care systems (e.g. admissions, billing, drug prescriptions), while research registries usually try to capture details on one specific disease or research question allowing for determinant-outcome like analysis. The use of medical registries for research purposes may thus be challenged by the limited amount of the information (including case mix), by the absence of determinant-outcome structure, and the limited quality. Quality concerns include: the classification of the health outcome, the limited amount of detailed clinical information (e.g. existing health status of the patient), the limited amount of information on events occurring before the health event (e.g. information on past exposures) and information on events after the health event (e.g. follow-up information).[27]

To evaluate the Dutch perinatal system, data from the Netherlands Perinatal Registry (PRN) were used. This registry reflects the complete Dutch perinatal experience from 2000-2007. Mortality data have been shown to be complete. No annual trends are observed between 2000-2007 in the relations shown, except for a small gradual decrease in total perinatal mortality.[2]

The PRN registry's quality is challenged in different ways. Health events and disease states are not always clearly classified, e.g. presence of pre-eclampsia or the start of delivery. Its quality is also challenged by the limited amount of information before and after birth (the health event), on former pregnancies, on risk factors (such as smoking, educational level, etc.), on process information of hospital admission and referrals (none of the interventions or diagnosis is combined with a date). Secondly, 68% of the paediatricians and 100% of NICU paediatricians participate. By the partial and selective participation of these paediatricians, completeness of short term neonatal outcome is challenged. Thirdly, the PRN does not contain long term follow up outcome of newborns. While these shortcomings may be true in general, our studies primarily suffer from the unavailability of detailed information on former pregnancies and risk factors, since these may lead to residual confounding. To some

extent they suffer from missing or confusing data on referrals and hospital admissions. Missing data on long term neonatal outcome made it impossible to compare long term effects (e.g. developmental disorders, psychopathological conditions, metabolic and cardiovascular disorders).

Despite these shortcomings, our studies show within its limits the invaluable availability of the PRN dataset for our research purposes.

Data obtained from the medical (administrative) registration of midwife practices to evaluate the birth centre Sophia contains information registered by the midwife only, its completeness is even more challenged as data of obstetrician or paediatrician or not included (sofar no joint electronic patient registry is available in the Netherlands).

The registry of the birth centre Sophia was intended to combine usefulness for care provision and simultaneous evaluation implying more research registry qualities. It includes more data on risk factors, includes more data and dates on process information (information on referrals, hospital admission), and uses clear classification of health outcomes. Yet, it still lacks data on long term neonatal outcome. The completeness and reliability of some variables were challenged since they are not yet routinely asked for during antenatal care. Practice research now addresses optimal recording of these data. Qualities of the ReproQ survey will be discussed in methodological issues part III.

Study design

A RCT would be the superior design to address our comparative research questions (I-V). It was difficult to pursue trials. When planned place of birth was part of a trial, it has been shown that participation hampered and introduced selective participation.[13,18-22,28] Another RCT on antenatal risk selection suffered from insufficient power.[11] Observational studies as ours are therefore of value, even if perfect adjustment or stratification can only be approximated.

Adjustment techniques

In our observational context different adjustment techniques are indispensible. We used multivariate regression techniques and (direct and indirect) standardization techniques. Major advantages of direct standardization compared to regression include computational simplicity and relatively few statistical assumptions. It is a preferred method when one is more interested in the overall effect instead of the influence of all the separate determinants. Standardized rates have been found to provide useful summary measures, especially when outcomes are

rare and specific rates display wide random variability. However, any summary measure can hide patterns that might have important public health implications. Standardization rates put more emphasis on less representative groups. Regression adjustment (forced entry), however, is more convenient for statistical tests for interactions and group differences (the individual effect of the different determinants). Additionally, when many determinants are present, regression is more convenient. However, more statistical assumptions have to be made in that case, such as; effects are linear correlated, effects are equal for everyone, and all combinations of parameters are possible (including biological implausible combinations).

Within standardization two methods are available. Direct standardization (weights taken from the index population) gives greater comparability but requires more data. Indirect standardization (weights taken from the standard population) requires fewer data but provides less comparability (unless the distribution of the standardization variable is identical across the study populations, in which case standardization is unnecessary since the crude death rates could have been compared directly). Decision is by data availability. Technically, therefore, one cannot compare indirect and direct standardization rates.

Big4/Big3 case mix adjustment

This thesis showed that the apparent contradiction on the safety of home deliveries in studies partly rests on a method choice.[29-37] Authors favouring a comparison of settings among 'suitable' home births only, usually exclude risk conditions were difference are observed between home and hospital births (Big3 pregnancies).

A second reason for the apparent contradiction in these studies rest on the degree of case mix adjustment. This thesis introduced the Big4/Big3 concept for additional case mix adjustment next to maternal factors, when primary data did not contain information on detailed risk factors. From detailed analysis of the complete perinatal dataset of the same Netherlands Perinatal Registry (PRN), years 2000-2007, (1.25 million records), it appeared that the presence of a Big4 condition preceded perinatal mortality in 85% of cases.[38] As Big4/Big3 adjustment therefore largely covers the patterns leading to perinatal mortality, and as the determination of Big4 conditions is not prone to much error, Big4/Big3 adjustment turns out to be a valuable concept. It can only be used for adjustment purposes, since it does not give information on the underlying risk factors leading to mortality, including information on their preventability. Overadjustment occurs when the occurrence of a Big4 condition and the pathway leading to mortality of this Big4 condition (case fatality) is influenced by a similar risk factor. Within our studies, overadjustment leads to underestimation of the true differences, as the direction of the effect of the risk factor is usually the same.

By Big4 adjustment we ignore differential management effects of setting during labour on the emergence of these Big4, should they exist (another example of overadjustment).

Exclusion of low Apgar from the Big4 (creating Big3) was done when evaluating the risk selection system and obstetric interventions, since these take place prior to the occurrence of low Apgar. However, when comparing mortality rates between the different birth settings we intentionally use Big4 as a risk indicator, because we assumed the role of management during labour to be small compared to the disadvantage of the home setting once a child with persistent low Apgar is born.

Methological issues PART III

Concept of responsiveness

First we discuss strengths and challenges of the Responsiveness concept in general. The eight domains were chosen based on a pre-existent philosophical structure, as identified in WHO's review of the patient satisfaction and quality of care literature, which also included the examination of different survey instruments.[39] Secondly, the independent value of the domains are supported by the human rights law which argues that the responsiveness features of a health system are important in their own right.[39-41] Thirdly, in contrast to patient satisfaction questionnaires, responsiveness tries to capture the patient's real experience, since literature has shown that expectations strongly influence patient satisfaction. Expectations may be influenced by economic and political influences and may lead to paradoxical results. For example, on going low economic resources on the country level or the personal level may lead to lower expectations and therefore a higher satisfaction, given that the quality of care remains the same. Expectations are also influenced by prior experiences and socio-demographic characteristics.[42-45] Lastly, aimed to develop a universal concept (e.g. developing and developed countries, different ethnicities, different care systems, etc.)[40]

The Responsiveness concept can be challenged by a number of issues. Firstly, capturing responsiveness by a limited number of questions with fixed answering categories may be a too restrictive approach. Combining qualitative research and different (quantitative) survey techniques, one can produce a richer dataset providing more explanatory information.[46] Secondly, the measurement of responsiveness even if it focsses on the patient's actual experience, is subject to elements of 'subjectivity'.

Thirdly, the concept of Responsiveness does not include financial barriers since these have to be evaluated according to the evaluation format by the WHO and according to standard practice in health care evaluation. Fourthly, the concept does not include technical quality

of professionals or setting. To the extent that technical quality skills improve health, it has to be captured by standard outcome measures such as perinatal mortality and morbidity in our case. Responsiveness captures the influence technical quality has on patient feelings. Lastly, the Responsiveness concept is a general concept for general health care. While attractive in general terms, the WHO concept should be checked whether it is sufficient to cover obstetric care completely. Our results show that the patient group of our studies felt all domains important, but this does not include lack of importance of some domains in specific groups. E.g. the domain social support appeared to be crucial for oncology patients[47], while this domain is much less appliccable to infertility patients.[48,49]

ReproQ

The ReproQ survey was offered post partum and addressed all the perinatal phases. Carry over effects on health responsiveness outcomes within the antenatal phase cannot be excluded, since birth outcome determinants significantly influenced outcomes within the antenatal phase. In the newly developed abridged digital form, we therefore separate an antenatal version from a postnatal version; questions are analogous but refer to different situations.

Although our study had a reasonable participation rate and covered all subgroups in Rotterdam, including groups which often refrain from participation in satisfaction surveys [50-51], no professional translator was used, but this was done by a family member of the women, introducing possible translation bias.

Barriers to the generalizability are the selection of the study population, whom primarily comes from a Dutch urban area, and the non-participation of women who did not understand the Dutch language sufficiently.

The ReproQ only focused on the responsiveness outcomes of the mother. Although, the domain 'Social Consideration' covered the role of the partner from the mothers view, the ReproQ did not include responsiveness outcomes of the partner.

RECOMMENDATIONS

From the above methodological and general discussion the following recommendations can be given.

Recommendations Part I and II

- A clear definition should exist on outcomes and risk factors when giving care and analysing care.
- Risk factors should be detected in a consistent way, preferably through a closed form checklist.

- ➢ Detailed risk factors should be recorded in a similar way and incorporated within the national perinatal registry.
- ➢ Medical (administrative) registries should be continually checked on quality, completeness and usefulness.
- ➢ A joint registry on pregnancy, mother and child outcomes should be developed by midwives, gynaecologists, youth health care providers and paediatricians under technical supervision of epidemiologists.
- ➢ Risk led care should be improved, in such a way that early interventions can be implemented based on risk factors leading to adverse mother and child outcome.
- ➢ To improve risk led care, non medical as well as medical factors should be taken into account and standardized protocols should be implemented.
- ➢ To improve on clinical and non clinical quality procedures should include measures to be effective in disadvantaged groups
- ➢ The expansion of midwife-led birth centres, with a similar organizational structure, should be stimulated; their use should be unrestricted by financial barriers
- ➢ The reintroduction of N_2O analgesia should be restricted to birth centres (adjacent to a hospital) and delivery rooms in a hospital.

Recommendations Part III

- ➢ Non clinical aspects should be evaluated using a WHO like tool.
- ➢ Non clinical aspects of antenatal care should be evaluated before delivery to prevent carry over effects of birth outcomes on the responsiveness outcomes of the antenatal phase.
- ➢ To improve non clinical outcomes of the Dutch Perinatal Care system in terms of responsiveness, caregivers should focus on domains covering the category ´client orientation´ (including the domains: Prompt Attention, Social Consideration, Quality of Basic Amenities, and Choice and Continuity).

REFERENCES

1. Mohangoo AD, Buitendijk SE, Hukkelhoven CW, et al. [Higher perinatal mortality in The Netherlands than in other European countries: the Peristat-II study]. Ned Tijdschr Geneeskd. 2008 Dec 13; 152(50):2718-27.
2. Ravelli AC, Tromp M, van Huis M, et al. Decreasing perinatal mortality in The Netherlands, 2000-2006: a record linkage study. J Epidemiol Community Health. 2009 Sep;63(9):761-5.
3. Zeitlin J, Wildman K, Breart G, et al. PERISTAT: indicators for monitoring and evaluating perinatal health in Europe. Eur J Public Health. 2003 Sep;13(3 Suppl):29-37.
4. de Graaf JP, Ravelli AC, Visser GH, et al. Increased adverse perinatal outcome of hospital delivery at night. BJOG. 2010 Aug;117(9):1098-107.

5. Evers AC, Brouwers HA, Nikkels PG, et al. Substandard care in delivery-related asphyxia among term infants: prospective cohort study. Acta Obstet Gynecol Scand. 2012 Sep 20.

6. Buitendijk SE, Nijhuis JG. [High perinatal mortality in the Netherlands compared to the rest of Europe]. Ned Tijdschr Geneeskd. 2004 Sep 18;148(38):1855-60.

7. Posthumus AG, Scholmerich VL, Waelput AJ, et al. Bridging Between Professionals in Perinatal Care: Towards Shared Care in The Netherlands. Matern Child Health J. 2012 Dec 11.

8. Halliday. J. WUDWAW 'Who Usually Delivers Whom and Where', Perinatal Data Collection Unit. Victoria, Australia: Victorian Government Department of Human Services. 1999.

9. Hampson JP, Roberts RI, Morgan DA. Shared care: a review of the literature. Fam Pract. 1996 Jun;13(3):264-79.

10. Tucker J, Florey CD, Howie P, McIlwaine G, Hall M. Is antenatal care apportioned according to obstetric risk? The Scottish antenatal care study. J Public Health Med. 1994 Mar;16(1):60-70.

11. Chan FY, Pun TC, Tse LY, Lai P, Ma HK. Shared antenatal care between family health services and hospital (consultant) services for low risk women. Asia Oceania J Obstet Gynaecol. 1993 Sep;19(3):291-8.

12. Perinatal and maternal outcomes by planned place of birth for healthy women with low risk pregnancies: the Birthplace in England national prospective cohort study. BMJ. 2011;343:d7400.

13. Eide BI, Nilsen AB, Rasmussen S. Births in two different delivery units in the same clinic—a prospective study of healthy primiparous women. BMC Pregnancy Childbirth. 2009;9:25.

14. Gottvall K, Waldenstrom U, Tingstig C, Grunewald C. In-hospital birth center with the same medical guidelines as standard care: a comparative study of obstetric interventions and outcomes. Birth. 2011 Jun;38(2):120-8.

15. Laws PJ, Tracy SK, Sullivan EA. Perinatal outcomes of women intending to give birth in birth centers in Australia. Birth. 2010 Mar;37(1):28-36.

16. Waldenstrom U, Nilsson CA. Characteristics of women choosing birth center care. Acta Obstet Gynecol Scand. 1993 Apr;72(3):181-8.

17. Gottvall K, Grunewald C, Waldenstrom U. Safety of birth centre care: perinatal mortality over a 10-year period. BJOG. 2004 Jan;111(1):71-8.

18. Hundley VA, Cruickshank FM, Lang GD, et al. Midwife managed delivery unit: a randomised controlled comparison with consultant led care. Bmj. 1994 Nov 26;309(6966):1400-4.

19. Waldenstrom U, Nilsson CA, Winbladh B. The Stockholm birth centre trial: maternal and infant outcome. Br J Obstet Gynaecol. 1997 Apr;104(4):410-8.

20. Begley C, Devane D, Clarke M, et al. Comparison of midwife-led and consultant-led care of healthy women at low risk of childbirth complications in the Republic of Ireland: a randomised trial. BMC Pregnancy Childbirth. 2011;11:85.

21. Bernitz S, Rolland R, Blix E, Jacobsen M, Sjoborg K, Oian P. Is the operative delivery rate in low-risk women dependent on the level of birth care? A randomised controlled trial. BJOG. 2011 Oct;118(11):1357-64.

22. Hodnett ED, Downe S., Edwards N., Walsh D. Home-like versus conventional institutional settings for birth. The Cochrane Collaboration. 2009.

23. Denktaş S, Waelput, A.J.M., van Voorst, S., Vos,A., Bonsel,G.J., de Jong-Potjer, L., Steegers, E.A.P. Healthy Pregnacy 4 All. Tijdschrift voor gezondheidswetenschappen. 2012.

24. Rosen MA. Nitrous oxide for relief of labor pain: a systematic review. Am J Obstet Gynecol. 2002 May;186(5 Suppl Nature):S110-26.

25. Bishop JT. Administration of nitrous oxide in labor: expanding the options for women. J Midwifery Womens Health. 2007 May-Jun;52(3):308-9.

26. Clyburn P. The use of Entonox for labour pain should be abandoned. Int J Obstet Anesth. 2001 Jan;10(1):27-9.

27. Reitsma J. Registers in cardiovascular epidemiology. Utrecht. 1999.

28. Hendrix M, Van Horck M, Moreta D, et al. Why women do not accept randomisation for place of birth: feasibility of a RCT in The Netherlands. Bjog. 2009 Mar;116(4):537-42; discussion 42-4.

29. Ackermann-Liebrich U, Voegeli T, Gunter-Witt K, et al. Home versus hospital deliveries: follow up study of matched pairs for procedures and outcome. Zurich Study Team. Bmj. 1996 Nov 23;313(7068):1313-8.

30. Anderson RE, Murphy PA. Outcomes of 11,788 planned home births attended by certified nurse-midwives. A retrospective descriptive study. J Nurse Midwifery. 1995 Nov-Dec;40(6):483-92.

31. Chamberlain G, Wraight A, Crowley P. Birth at home. Pract Midwife. 1999 Jul-Aug;2(7):35-9.

32. de Jonge A, van der Goes BY, Ravelli AC, et al. Perinatal mortality and morbidity in a nationwide cohort of 529,688 low-risk planned home and hospital births. Bjog. 2009 Aug;116(9):1177-84.

33. Janssen PA, Saxell L, Page LA, Klein MC, Liston RM, Lee SK. Outcomes of planned home birth with registered midwife versus planned hospital birth with midwife or physician. Cmaj. 2009 Sep 15;181(6-7):377-83.

34. Lindgren HE, Radestad IJ, Christensson K, Hildingsson IM. Outcome of planned home births compared to hospital births in Sweden between 1992 and 2004. A population-based register study. Acta Obstet Gynecol Scand. 2008;87(7):751-9.

35. Pang JW, Heffelfinger JD, Huang GJ, Benedetti TJ, Weiss NS. Outcomes of planned home births in Washington State: 1989-1996. Obstet Gynecol. 2002 Aug;100(2):253-9.

36. Wax JR, Pinette MG, Cartin A, Blackstone J. Maternal and newborn morbidity by birth facility among selected United States 2006 low-risk births. Am J Obstet Gynecol. 2010 Feb;202(2):152 e1-5.

37. Wiegers TA, Keirse MJ, van der Zee J, Berghs GA. Outcome of planned home and planned hospital births in low risk pregnancies: prospective study in midwifery practices in The Netherlands. Bmj. 1996 Nov 23;313(7068):1309-13.

38. Bonsel GJ, Birnie E, Denktas S, Poeran J, Steegers EAP. Dutch report: Lijnen in de Perinatale Sterfte, Signalementstudie Zwangerschap en Geboorte 2010 (http://www.nvk.nl/Nieuws/Dossiers/DossierPerinatalezorg.aspx). Rotterdam: Erasmus MC; 2010.

39. de Silva A (ed). A framework for measuring responsiveness. Global Programme on Evidence for Health Policy Discussion Paper. No. 32. Geneva: World Health Organisation.

40. Valentine NB, de Silva A, Kawabata K, Darby C, Murray CJL, Evans DB (eds). Health system responsiveness: concepts, domains and measurement. Geneva: World Health Organisation; 2003.

41. Gostin L (ed). The domains of health responsiveness: a human rights assessment. Geneva: World Health Organisation; 2002.

42. Murray CJ, Kawabata K, Valentine N. People's experience versus people's expectations. Health Aff (Millwood). 2001 May-Jun;20(3):21-4.

43. Hall JA, Irish JT, Roter DL, Ehrlich CM, Miller LH. Satisfaction, gender, and communication in medical visits. Med Care. 1994 Dec;32(12):1216-31.

44. Murphy-Cullen CL, Larsen LC. Interaction between the socio-demographic variables of physicians and their patients: its impact upon patient satisfaction. Soc Sci Med. 1984;19(2):163-6.

45. Blendon RJ, Kim M, Benson JM. The public versus the World Health Organization on health system performance. Health Aff (Millwood). 2001 May-Jun;20(3):10-20.

46. Foundation NS. User friendly Handbook for mixed evaluations. 1997.

47. Ouwens M, Hermens R, Hulscher M, et al. Development of indicators for patient-centred cancer care. Support Care Cancer. 2010 Jan;18(1):121-30.

48. Whiteford LM, Gonzalez L. Stigma: the hidden burden of infertility. Soc Sci Med. 1995 Jan;40(1):27-36.

49. van Empel IW, Dancet EA, Koolman XH, et al. Physicians underestimate the importance of patient-centredness to patients: a discrete choice experiment in fertility care. Hum Reprod. 2011 Mar;26(3):584-93.

50. Gayet-Ageron A, Agoritsas T, Schiesari L, Kolly V, Perneger TV. Barriers to participation in a patient satisfaction survey: who are we missing? PLoS One. 2011;6(10):e26852.

51. Boshuizen HC, Viet AL, Picavet HS, Botterweck A, van Loon AJ. Non-response in a survey of cardiovascular risk factors in the Dutch population: determinants and resulting biases. Public Health. 2006 Apr;120(4):297-308.

13

Summary/samenvatting

SUMMARY

The assumed merits of the Dutch Perinatal System have come under scrutiny since the national perinatal mortality rate, defined as mortality between 22 weeks of pregnancy and one month after delivery, consistently showed to be one of the highest in Western-Europe.

The overall aim of this thesis was: (1) to give possible explanations for the high perinatal mortality in the Netherlands, in particular explanations related to the Dutch Perinatal Care system; (2) to evaluate potential innovations to improve outcome of the Dutch Perinatal Care system; and (3) to introduce new evaluation methods suitable to detect changes induced by innovations of the Dutch Perinatal Care system.

All new evaluation methods used the concept of Big3/Big4, a concept to adjust for perinatal risk differences, and the concept of responsiveness, introduced by the World Health Organization (WHO) to evaluate the non clinical aspects of care.

Big3 pregnancies are defined as: congenital abnormalities (list defined), intrauterine growth restriction (Small for Gestational Age, birth weight below the 10th percentile for gestational age, gender and parity specific), or preterm birth (< 37th week of gestation). Big4 pregnancies additionally include low Apgar score (< 7, measured 5 minutes after birth). The concept of Big3/Big4 is part of the two stage epidemiological model used in this thesis to describe and explain perinatal morbidity. The first stage is the antenatal development of such a morbidity, the second stage, eventually, is the subsequent fetal or neonatal death. This concept is justified by the observation that 85% of current perinatal mortality is associated with Big4 morbidity. The Big3/Big4 concept appeared fruitful in case mix adjustment procedures and in the distinction between opportunities for better performance at the early antenatal stage versus those prior to and during delivery.

The concept of responsiveness as introduced by the WHO reflects the client orientation of care providers as perceived through the eyes of the client. The WHO was responsible for introducing this concept at a global scale, after lengthy consultation of a broad range of stakeholders. It showed suitable for health system evaluation at large as well as for evaluation of specific health system services. The concept can be considered universal, and we derived a perinatal elaboration sticking close to the original intent.

In **chapter 1** briefly introduces the research described in this thesis.

Part I (chapter 2, 3 and 4) elaborates on the possible explanations for the high perinatal mortality in the Netherland, with an emphasis on the performance of the Dutch Perinatal Care system rather than on socio-demographic factors. These chapters use clinical outcomes

to measure quality of care. For all analyses data from the Dutch Perinatal Registry (PRN) was used. This registry reflects the complete Dutch perinatal care experience from 2000-2007.

Chapter 2 elaborates on the effectiveness of the current risk selection within the Dutch Perinatal system. For that purpose we established the degree to which high risk (Big3) pregnancies were referred prior, during or after birth, where referral prior to birth is the preferred outcome. Selection failure is approximated by the degree to which Big3 conditions are present at the start of midwife-led deliveries, and the amount of late referral (for those reasons) at parturition. Our analysis shows that the current risk selection occasionally classifies as low risk pregnancies where risk is increased. **Chapter 3** elaborates on perinatal outcomes comparing planned home births with planned hospital births, both mid-wife led births. Data showed that planned home birth, under routine conditions, generally is not associated with increased intrapartum and early neonatal death, yet in unnotified Big4 pregnancies additional risk cannot be excluded. **Chapter 4** compares the intervention rate between planned home versus planned hospital births; it determines, if differences in intervention rates exist, whether these could be interpreted as over- or undertreatment by comparing perinatal mortality rates taking care for case mix. We observed, as other did before, that the planned place of birth has an impact on the intervention rate. Multiparous women showed universally lower intervention rates in planned home births, while primiparous women showed rather similar intervention rates. The presence of over- or undertreatment expressed by perinatal mortality differs per risk group. In this assumed low risk population, undertreatment in the home setting was observed in undetected risk groups using the Big3/Big4 concept.

Our results give opportunities for the Dutch perinatal system to improve. Three innovative strategies to improve perinatal care were implemented and evaluated, covering the structural, process and outcome level respectively (see introduction).

Part II (chapter 5, 6, 7, and 8) elaborates on possible innovative strategies to improve risk selection and midwife-led birth care. To evaluate the innovative strategies of the midwife-led birth centre on clinical outcomes and risk led care, data was obtained from the medical registration of midwife practices and from the administrative registration of the birth centre Sophia adjacent to the Erasmus Medical Centre.

Chapter 5 evaluates the introduction of this midwife-led birth centre, which adopted an innovative strategy focussed on advanced risk management and integral care delivery. Integral means that different caregivers truly collaborate, use the same risk concepts, and thoroughly communicate on risks and their approach during the stay at the birth centre and thereafter. Evaluation was done by comparing population characteristics, adverse perinatal and maternal outcomes, and intervention rates between planned home births,

planned hospital births and planned birth centre births, all supervised by the community midwife. The introduction of a midwife-led birth centre led to a redistribution of planned place of midwife-led births. The women planning their delivery in the midwife-led birth centre had the highest risk profile compared to women planning their delivery at home or in the hospital. After the introduction of the midwife-led birth centre, a decreasing trend of neonatal mortality and morbidity, and maternal morbidity was observed in planned home births, hospital births and birth centre births, suggests on average better care through more adequate selection of the best setting. Overall intervention rates appeared unaffected by the introduction of the birth centre. **Chapter 6** evaluates a second innovation, namely the introduction the Rotterdam Reproductive Risk Reduction (R4U). The R4U is developed to improve risk selection using medical as well as non medical risk factors. The R4U was implemented and used to improve the detection of Small for Gestational Age (SGA) at the start of birth. We showed that the R4U, after including ethnicity, obstetric and social risk factors still detects SGA cases which were unnoticed antenatally. We therefore consider the R4U a valuable adjuvant diagnostic tool for this purpose (predelivery risk assessment). **Chapter 7 and 8** evaluates the third innovation, namely the introduction of nitrous oxide analgesia during labour in the midwife-led birth centre Sophia. The occupational safety and effectiveness of nitrous oxide analgesia were evaluated. While administrating nitrous oxide during labour in the midwife-led birth centre Sophia, occupational exposure limits were met. It relied on the use of a strict protocol and the standardized use of the Anevac P-scavenging system. The introduction of nitrous oxide led to non trivial substitution of pethidine and/or epidural analgesia (5%). Women receiving nitrous oxide showed significant decrease in feeling nervous, anxious and pain. The availability of nitrous oxide analgesia led to an additional increase of analgesia use in general (8%) compared to the period when nitrous oxide was not available.

Part III (chapter 9, 10 and 11) elaborates on a new evaluation method to evaluate the performance of the Dutch Perinatal Care system on non clinical aspects of care quality. It adopted the responsiveness concept of the World Health Organization. Responsiveness is defined as the way an individual is treated and the environment in which she is treated during health system interactions. We derived the ReproQ questionnaire for use in perinatal care. This questionnaire translates the eight domain responsiveness questionnaire developed by the World Health Organization (WHO) into the perinatal care context. The eight domains are: dignity, autonomy, respect, communication, prompt attention, basic amenities, social support, and choice and continuity; each domain is

covered by a small set of questions on the client's experience, in our case antenatally and during delivery.

Chapter 9 showed the psychometric properties (feasibility, validity) of the newly developed responsiveness questionnaire ReproQ. Women who delivered in either in primary, secondary or tertiary care were asked to participate. Participation included an interview 2-8 weeks post partum. The ReproQ questionnaire demonstrated high acceptation of low socio-economic and Dutch speaking ethnic groups. It also showed satisfactory psychometric properties. We conclude the ReproQ has the potential to discriminate responsiveness across different experiences, but that routine application in the absence of interviewer's support requires an abridged version.

Chapter 10 explored whether different domain weights should be used for subgroups within the Dutch Perinatal System. The importance of the domains was assessed for the subgroups. It showed that individual factors had no significant effect on the relative importance of domains and items. Therefore, all perinatal subpopulations can use the same set of equally valued quality ('responsiveness') domains and items, a criticial feature.

Chapter 11 explored responsiveness results of the Dutch Perinatal Care system. Generally, responsiveness outcomes of the Dutch Perinatal Care system performs quite well in absolute terms. Responsiveness of the four 'client orientations' domains (prompt attention, basic amenities, social support, choice and continuity) underperformed compared to the 'respect to persons' domains (dignity, autonomy, respect, communication). Background characteristics show little systematic impact on responsiveness, a desired measurement property.

Chapter 12, the general discussion, combines the results of all studies and discusses these from a broader perspective. Our stated goals were to find possible explanations for the high perinatal mortality in the Netherlands in relation to the performance of the Dutch Perinatal Care, to provide give some innovations to improve on this performance, and to evaluate the Dutch system on non clinical outcomes additional to the conventional clinical ones.

Underperformance of the Dutch Perinatal Care system could only partially evaluated, and in some respects provided an explanation for its underperformance. Other explanations are selected maternal factors, organizational factors, and policy factors. In this thesis we concluded that especially in the area of risk management considerable improvements should be made and one may expect these improvements to improve the performance of the Dutch Perinatal system at a larger scale. The innovations can be at the system, the professional level, and to a lesser extent, the client level.

SAMENVATTING

Vergeleken met andere West - Europese landen heeft Nederland een hoge perinatale sterfte, gedefinieerd als sterfte vanaf 22 weken zwangerschap tot en met één maand na de bevalling. Hierdoor zijn vraagtekens gezet bij het functioneren van het huidige Nederlands Verloskundig Systeem.

Dit proefschrift heeft als doel om: (1) mogelijke verklaringen te vinden voor de hoge perinatale sterfte in Nederland, in het bijzonder in het functioneren van het Nederlands Verloskundig Systeem; (2) innovaties ter verbetering van het Nederlands Verloskundige systeem te evalueren; (3) om voor deze innovaties in het Nederlands Verloskundig Systeem nieuwe evaluatiemethoden te introduceren.

Deze nieuwe evaluatiemethoden zijn zowel gebaseerd op het concept Big3/Big4, waarbij verschillen in case mix tussen groepen zoveel mogelijk gelijk getrokken worden, als op het World Health Organization (WHO) concept responsiveness concept, om de niet-klinische aspecten van zorg te evalueren.

De definitie van een Big3 zwangerschap is: congenitale afwijking, groeivertraging ('Small for Gestational Age'; te klein voor de zwangerschapsduur met een geboortegewicht onder het 10de percentiel), en premature geboorte (<37 weken zwangerschapsduur). De definitie van een Big4 zwangerschap voegt daar de score van een lage Apgar score (<7 na 5 min) bij.

Het Big3/Big4 concept maakt onderscheid tussen twee verschillende fasen die vooraf gaan aan perinatale sterfte, namelijk de fase die leidt tot het ontstaan van een Big3/Big4 conditie en de fase waarin een Big3/Big4 conditie bestaat en deze leidt tot perinatale sterfte.

Het WHO concept responsiveness evalueert de patiëntgerichtheid van een zorgsysteem zoals ervaren wordt door de patiënt op niet klinische uitkomsten, zoals bijvoorbeeld de wachttijd tot zorg. De WHO introduceerde dit concept om op deze manier zowel zorgsystemen op grote schaal als specifieke zorgsystemen te kunnen evalueren, ongeacht de precieze kenmerken van het systeem of zorgverlener. Het kan daarom als universeel beschouwd worden.

In **hoofdstuk 1** geven we een inleiding op het onderzoek dat in dit proefschrift wordt gepresenteerd.

Deel I (hoofdstuk 2, 3 en 4) beschrijft de studies met de mogelijke verklaringen voor de hoge perinatale sterfte in Nederland in relatie tot het functioneren van het Nederlands Verloskundig Systeem. Hiervoor zijn klinische uitkomstmaten gebruikt uit de Perinatale

Registratie Nederland (PRN). Deze nationale registratie bevat informatie over nagenoeg alle bevallingen in Nederland tussen 2000-2007.

Hoofdstuk 2 beschrijft de effectiviteit van de huidige manier van risicoselectie binnen het Nederlands Verloskundig Systeem. Om de effectiviteit van risicoselectie binnen het Nederlands verloskundig systeem te evalueren, kijken we naar het aantal verwijzingen van hoog risico (Big3) zwangerschappen voor, tijdens of na de bevalling. Onze resultaten lieten zien dat bij de huidige risicoselectie een klein aantal zwangerschappen onterecht als laagrisico zwangerschap werd gekwalificeerd, waardoor de start bij bevalling in de eerstelijn plaats vindt, terwijl die eigenlijk in de tweedelijn had moeten plaatsvinden. In **Hoofdstuk 3** wordt de perinatale sterfte bij alle geplande thuisbevallingen die hun baring starten in de eerstelijn vergeleken met alle geplande poliklinische bevallingen die hun baring starten in de eerstelijn. Resultaten laten zien dat thuisbevallingen over het algemeen niet geassocieerd zijn met een hogere perinatale sterfte. Een verhoogd risico voor niet opgespoorde Big4 zwangerschappen kon bij een geplande thuisbevalling echter niet worden uitgesloten.

Hoofdstuk 4 vergelijkt het aantal medische interventies (kunstverlossing en/of keizersnee) tussen geplande thuisbevallingen en geplande poliklinische bevallingen onder leiding van een verloskundige. In dit hoofdstuk stellen we vast, met behulp van perinatale sterftecijfers, of verschillen verklaart kunnen worden door over- of onderbehandeling. Het percentage medische interventies in geplande thuisbevallingen was in multipara vrouwen over het algemeen lager dan in geplande poliklinische bevallingen, terwijl het percentage interventies in primipara vrouwen nagenoeg niet verschilde. De mogelijke aanwezigheid van over- of onderbehandeling verschilt per risicogroep. In Big3 zwangerschappen was de perinatale sterfte bij geplande thuisbevallingen hoger dan bij geplande poliklinische bevallingen en onderbehandeling mogelijk aanwezig.

De resultaten zoals hierboven beschreven zijn, bieden mogelijkheden om het Nederlands Verloskundig Systeem te verbeteren. Drie innovatieve strategieën om perinatale zorg te verbeteren werden geïmplementeerd en geëvalueerd.

In **Deel II** (hoofdstuk 5, 6, 7 en 8) worden mogelijke strategieën om risicoselectie en baringszorg onder leiding van de verloskundige te verbeteren en het functioneren van het Nederlands Verloskundige systeem te beïnvloeden beschreven en geëvalueerd. Om de verbeteringsstrategieën te evalueren, is gebruik gemaakt van de medische registratie van de lokale verloskundige praktijken en van de medische registratie van het geboortecentrum Sophia.

In **Hoofdstuk 5** wordt de introductie van het eerstelijns geboortecentrum Sophia als zorg innovatie beschreven, en de evaluatie daarvan. Centraal staan de gerichtheid op risico geleide zorg en het streven naar verbeterde ketenzorg. Vergeleken worden de populatiekarakteristieken, ongunstige kind- en moeder uitkomsten en het aantal interventies tussen geplande thuisbevallingen, poliklinische bevallingen en bevallingen in het geboortecentrum. Vrouwen die hun bevalling in het eerstelijns geboortecentrum planden, hadden een hoger risicoprofiel vergeleken met vrouwen die hun bevalling thuis of poliklinisch planden. In de periode nadat het geboortecentrum is geopend, werd een daling gezien van ongunstige kind- en moeder uitkomsten. Deze daling lijkt te komen door een betere risicoselectie voor de geplande plaats van de bevalling. Het percentage medische interventies in deze twee perioden bleef gelijk.

In **Hoofdstuk 6** staat een tweede zorginnovatie centraal, namelijk het instrument de Rotterdam Reproductive Risk Reduction (R4U). De R4U is ontwikkeld om de risicoselectie te verbeteren en maakt gebruik van medisch en niet medische risico factoren. De R4U werd ingevoerd en getest om de aanwezigheid van eerder (gemiste) Small for Gestational Age (SGA) op te sporen bij de start van bevalling. De R4U was in staat om gemiste SGA op te sporen. Opsporen van gemiste SGA werd verbeterd wanneer er gebruik werd gemaakt van zowel medische (algemene en obstetrische) als niet-medische (psycho-sociale en maatschappelijke) risico factoren. We beschouwen de R4U daarom als een waardevol aanvullend diagnostisch instrument (risico inschatting) voor de bevalling.

De derde zorginnovatie komt ter sprake in **Hoofdstuk 7 en 8**. Hier ging het om de introductie van lachgas analgesia tijdens de bevalling in het eerstelijns geboortecentrum Sophia. De uitkomsten van deze vorm van pijnstilling werden geëvalueerd in termen van werkveiligheid voor de zorgverlener en de effectiviteit voor de patiënt. Binnen het geboortecentrum Sophia werd lachgas volgens de huidige ARBO richtlijnen toegediend en kreeg zij officieel toestemming van de arbeidsinspectie voor het toedienen van lachgas. Dit werd bereikt door strikt te werken volgens gestandaardiseerde protocollen en systematisch gebruik te maken van het Anevac P afzuigsysteem. De mogelijkheid om te bevallen met lachgas heeft ervoor gezorgd dat, met name multipara vrouwen, lachgas analgesie kregen in plaats van pethidine en/of epiduraal (5%). De vrouwen die lachgas kregen lieten een daling zien in gespannenheid, angst en pijnscores. Er vond wel een toename van pijnstilling verzoek plaats (8%) ten opzichte van de jaren waarin lachgas niet tot een van de opties behoorden.

Deel III (hoofdstuk 9, 10, en 11) gaat over de evaluatie van het Nederlands Verloskundig systeem in termen van niet-klinische uitkomsten. Het beschrijft een nieuwe methode om niet-klinische uitkomsten van het Nederlands Verloskundig Systeem te evalueren, volgens

het World Health Organization (WHO) concept responsiveness. Responsiveness geeft een oordeel over de manier waarop patiënten worden behandeld en in welke omgeving waarin zij worden behandeld. Om responsiveness te kunnen meten werd het ReproQ instrument ontwikkeld. Deze werd ontwikkeld door de responsiveness vragenlijst van de WHO, bestaande uit acht domeinen, te vertalen en zo nodig op verloskundige zorg aan te passen. De acht domeinen zijn: respect, autonomie, privacy, communicatie, tijd tot zorg, kwaliteit faciliteit, sociale steun, en keuze en continuïteit zorgverlener.

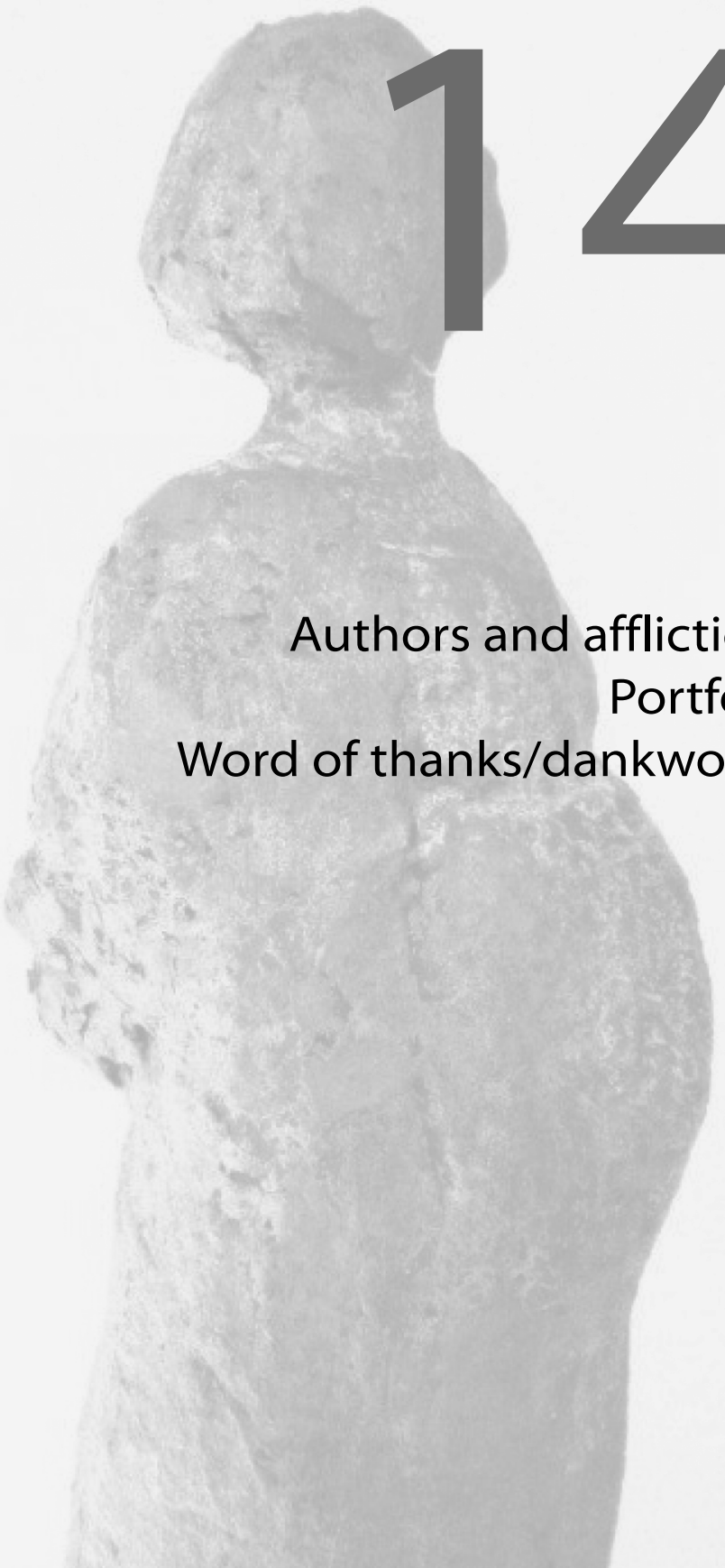
In **Hoofdstuk 9** worden de psychometrische kenmerken (hanteerbaarheid, validiteit) van de nieuw ontworpen responsiveness vragenlijst ReproQ beschreven. Vrouwen die bevallen waren in de eerste, tweede of derde lijn werden gevraagd deel te nemen. Deelname bestond uit een interview dat 2-8 weken post partum werd gehouden. De ReproQ laat goede psychometrisch eigenschappen zien en kan discrimineren tussen responsiveness uitkomsten van verschillende ervaringen.

In **Hoofdstuk 10** wordt onderzocht of domeinscores zomaar bij elkaar opgeteld mogen worden of dat er rekening gehouden moet worden met de belangrijkheid van domeinen voor verschillende subgroepen van patiënten. Onderzocht wordt of domein gewichten (belangrijkheid) verschillend zijn voor subgroepen van patiënten binnen de verloskundige zorg. We vonden dat individuele factoren geen significant effect hebben op domein gewichten. Wanneer responsiveness wordt gemeten in de verloskundige zorg is het niet nodig de verschillende domeinen te wegen voor subgroepen.

Hoofdstuk 11 geeft een beschrijving van de uitkomsten van responsiveness van het Nederlands Verloskundig Systeem. De responsiveness uitkomsten zijn over het algemeen goed. Voor de vier 'cliënt georiënteerde' domeinen (tijd tot zorg, kwaliteit faciliteit, sociaal support, en keuze en continuïteit zorgverlener) werden slechtere responsiveness uitkomsten gevonden dan voor de 'respect voor personen' domeinen (respect, autonomie, privacy, communicatie). Karakteristieken van vrouwen hadden weinig systematisch effect op responsiveness uitkomsten. Zorguitkomsten hadden een effect op met name de responsiveness uitkomsten van de baring.

In het laatste hoofdstuk, **hoofdstuk 12**, worden alle resultaten van de bovenstaande studies in een breder perspectief gezet en kritisch besproken. Het doel was om mogelijke verklaringen te vinden voor de hoge perinatale sterfte en het niet goed functioneren van het Nederlands Verloskundige Systeem. Daarnaast het ontwikkelen en evalueren van innovaties ter verbetering van het Nederlands Verloskundige Systeem, en hiervoor nieuwe evaluatiemethoden te introduceren.

Het huidig functioneren van het Nederlands Verloskundig systeem wordt op een aantal punten geëvalueerd. De beschreven studies geven ten delen een verklaring voor de hoge perinatale sterfte in Nederland. Dit proefschrift lieten we zien dat een mogelijke verklaring van de hoge perinatale sterfte met name gezocht moet worden in het gebied van risico selectie en management. Andere verklaringen moeten worden gezocht in organisatie en registratie van zorg. In dit proefschrift lieten we enkele veelbelovende strategieën voor verbeteringen zien op zowel het systeem als professioneel niveau.



14

Authors and afflictions
Portfolio
Word of thanks/dankwoord

AUTHORS AND AFFLICTIONS

From the Department of Obstetrics and Gynaecology, Division of Obstetrics and Prenatal Medicine, Erasmus MC, University Medical Centre Rotterdam, The Netherlands

Semiha Denktaş, Eric A.P. Steegers, Gouke J. Bonsel, Hanneke P. de Graaf, Johannes J. Duvekot, Elaine Fitzpatrick, Laura B.E.S. Quispel, Myrna Kastaneer, Kim van der Lede-Fabricius.

From the Department of Anaestesiology, Erasmus MC, University Medical Centre Rotterdam, The Netherlands

Ilse J.J. Dons

From the Department of Public Health, Erasmus MC, University Medical Centre Rotterdam, The Netherlands

Gouke J. Bonsel

From the Midwifery Academy, Rotterdam University of Applied Sciences, Rotterdam, The Netherlands

Gouke J. Bonsel

From the Institute of Health Policy and Management, Erasmus University, Rotterdam, The Netherlands

Erwin Birnie

From the World Health Organization, Social Determinants of Health - Ethics, Equity, Trade and Human Rights - Innovation, Information, Evidence and Research, Avenue Appia 20, 1211 Geneva, Switzerland.

Nicole B. Valentine

From the Department of Health, Safety and Environment, Erasmus MC, University Medical Centre Rotterdam, The Netherlands

Zita M. Kolder, Karel D. Witters

PORTFOLIO

Summery of PhD training and teaching

Name PhD candidate	Jacoba van der Kooy
Erasmus MC Department	Obstetrics and Gynaecology Division of Obstetrics and Prenatal Medicine
Research School	Netherlands Institute for Health Science (NIHES)
PhD period	August 2009 – December 2012
Promotors	Prof. dr. Eric A.P. Steegers Prof. dr. Gouke J. Bonsel
Copromotors	Dr. Semiha Denktaş Dr. Erwin Birnie

General courses	Year	Workload (ECTS)
Master of Science in Clinical Epidemiology, NIHES, Rotterdam, The Netherlands	2008	35
Public Health Studies, Johns Hopkins University, Baltimore, United States	2008	5
Master of Science in Epidemiology, NIHES, Rotterdam, The Netherlands	2012	35
International and National conferences		
Society for Gynecologic Investigation (SGI) congress, Orlando. Poster presentation.	2010	1
Verloskundig Samenwerkingsverband (VSV) Schiedam, The Netherlands. Oral presentation.	2010	1.5
Symposium 'Pijn tijdens de bevalling', Utrecht, The Netherlands. Oral presentation.	2010	1.5
Society for Gynecologic Investigation (SGI) congress, Miami. Poster presentation.	2011	1
Grootstedelijke perinatale gezondheid congres, Rotterdam, The Netherlands. Oral presentation.	2011	1.5
Symposium 'Bevallen, thuis of poliklinisch?', Maastricht, The Netherlands. Oral presentation.	2011	1.5
Society for Gynecologic Investigation (SGI) congress, San Diego. Poster presentation.	2012	1
First International Conference on Fetal Growth, Birmingham, United Kingdom. Oral presentation.	2012	1.5
European Public Health Conference (EPH), St. Julian's Bay, Malta. Oral presentation.	2012	1.5
Ikazia 'Lachgas tijdens de bevalling', Rotterdam, The Netherlands. Oral presentation.	2012	1.5
Symposium 'Geboortecentrum 2,5 jaar', Rotterdam, The Netherlands. Oral presentation.	2012	1.5

Other		
Grant Proposal assistance (ZonMW)	2009/ 2012	5
Local meetings		
Kring symposium, Rotterdam, The Netherlands.	2009	
Kring symposium, Rotterdam, The Netherlands.	2010	
Café BéBé meetings	2010/ 2012	
Research meetings	2009/ 2012	
Teaching activities		
Geneeskunde minor, Cylcle of Life, Rotterdam, The Netherlands.	2010	5
Opleiding Klinisch verloskundige, a Midwife-led Birth Centre, Rotterdam, The Netherlands.	2011	2
NIHES, Perinatal care in Urban areas; a midwife-led birth centre, Rotterdam, The Netherlands.	2012	1
Supervising students		
Laura Quispel, Student Geneeskunde, Erasmus MC.	2010	2
Projectmatig en Multidiciplinair werken aan Grootstedelijke Vraagstukken (PMG).	2010/ 2011	10
(Marit Naerebout, Janine Goldmann, Mylene van de Voort Wietske Stout, Laurentine van Wessem, Mandy van der Heiden Annewil van Veen, Inez van Vlijmen, Arianne van Houwelingen Juanite Watson, Marije van der Werf, Birgit Dijkstra Corine de Lange), Studenten zorgsector, Hogeschool Rotterdam.		
Suzanne Ammerlaan, Student Verloskunde, Verloskundige Academie Rotterdam.	2011	2
Nathalie Boogers, Student Verloskunde, Verloskundige Academie Rotterdam.	2011	2
Kim fabricius, Student Geneeskunde, Erasmus MC.	2011	2
Myrna Kasteneer, Student Geneeskunde, Erasmus MC.	2011	2
Anniek Feijge, Student Pedagogiek, Hogeschool Rotterdam.	2011	2
Jet van der Kooy, Student Pedagogiek, Hogeschool Rotterdam.	2011	2

WORD OF THANKS/DANKWOORD

Na bijna vier jaar onderzoek met SPSS, Excel, analytische beschouwingen, syntax, imputeren, strakke teksten, Ctrl sneltoetsen, track changes, powerpoint, pubmed MESH terms en 'submission approved' is het dan zo ver: mijn proefschrift is klaar. Alleen nog het dankwoord. Dat blijkt een terugblik vol goede herinneringen. Naast het onderzoekswerk, was het deel uitmaken van een maatschappelijk multicultureel dynamisch project, gericht op verbetering van verloskundige zorg in de stad Rotterdam, een ervaring die mij nog lang zal bijblijven.

Zonder hulp van anderen was dit proefschrift er niet gekomen. Er zijn veel mensen die mij direct of indirect hebben geholpen dit proefschrift te voltooien. Een aantal van hen wil ik hier persoonlijk bedanken.

Professor Eric Steegers, beste Eric. Ik herinner mij nog goed de presentatie die u gaf op de Bella Obstetrica over het programma 'Klaar voor een Kind'. Daar wilde ik graag aan meewerken en die kans heb ik gekregen, bedankt! Ooit vroeg ik u wat u dreef. U antwoordde toen; 'Leuke dingen blijven doen'. En dat hebben we gedaan. Ik heb zeer genoten van de samenwerking in uw onderzoeksgroep en ik hoop in de toekomst nog 'leuke dingen' te blijven doen.

Professor Gouke Bonsel, Gouke waar zal ik beginnen. Wat heb ik veel van je geleerd: systematisch denken, analyseren, Excel smart functies, schrijven, precisie. Er was geen onderwerp waar jij al niet wat over had gelezen of onderzoek naar had gedaan. Ik bewonder je creatieve geest en passie voor het werk. Het was niet ongewoon dat je met een vraag 'even' binnenkwam en met drie onderzoeksvoorstellen en een persoonlijkheidsanalyse de deur uit ging. Ja, ik geef toe dat het soms niet makkelijk was met dit ongeduldige, soms eigenwijze meisje. Bedankt.

Ja en zonder Gouke geen Erwin, de altijd stabiele, onverstoorbare en geduldige Erwin met zijn droge humor. Erwin, bedankt dat je altijd klaar stond als ik syntax problemen had, voor je geduld om mijn stukken weer door te kijken en dat je altijd de dingen heerlijk wist te relativeren.

Semiha, ik bewonder je gedrevenheid, je oog voor kansen en je positieve denken. Ik heb erg genoten van je humor, je luisterend oor, de scherpe analyses over de Nederlandse cultuur en je liefde voor salades.

Hanneke, ´blijven doorgaan totdat het voor elkaar is´ dat is wat jou op het lijf is geschreven. Met gepaste trots kunnen we terug zien op de totstandkoming van een prachtig geboortecentrum. Bedankt voor de vrijheid en het vertrouwen om dingen te ontwikkelen en te implementeren. En natuurlijk zal ik nooit die schaterende en galmende lach vergeten.

Kraamverzorgenden en verloskundigen van het geboortecentrum. Zonder jullie was dit boekje er niet gekomen. Bedankt dat jullie altijd weer open stonden voor vernieuwingen en meedachten over verbeteringen. En ja, daar kwam ik weer met mijn lijstjes. Expertisegroep, ik heb genoten van de maandelijkse brainstormsessies, de scherpheid en natuurlijk de gezelligheid.

Fellow onderzoekers, het hoogtepunt was toch wel de woensdagochtend research bespreking. Een stukje wetenschap en gezelligheid. Deze combi kwam nog sterker terug op de congressen waar we erg genoten hebben. Research, lief en leed werden daar gedeeld. Als ik over congressen begin dan denk ik gelijk aan Fatima, Marijana, en Jashvant. Een gemêleerder stel is niet voor te stellen. Vele mensen hebben zich afgevraagd hoe deze Marokkaanse moslima, orthodoxe Balkanese , Hindoestaanse Surinamer en christelijke Hollandse het zo goed met elkaar kunnen vinden. Wat heb ik genoten van jullie humor, van de hart tot hart gesprekken, de weekenden weg, de stress op Charleroi, de vichy van Jashvant, de 'large cappachino with soy milk' en een extra shot espresso, ontbijt op het plein van de Sint Pieter, 'ons hotel' in Orlando, Warung Mini, 3kg Turks fruit in Istanbul… 'We will meet in New York'.

Jashvant, daar staan we dan. Wat een herinneringen. We begonnen staand bij de kopieermachine: het eerste R4U partusverslag 'founded by'. Vanaf toen waren we een team. Het is moeilijk in woorden op te schrijven hoeveel ik geniet van onze vriendschap, van jouw humor, je relativeringsvermogen, je oprechte interesse in mensen, het subtiel voorhouden van een spiegel en je gave om echt te kunnen genieten. You always got my back, and as you say: 'You are a true friend!' And remember, we were there, before and after the 'signalement'.

Dit dankwoord schrijven zet mij stil bij de mensen om mij heen. Wat een rijkdom. Lieve Dien, er zijn maar weinig mensen die mij zo goed kennen als jij. Bij het opnemen van de telefoon weet jij al hoe laat het is. Bewegen, koffie, kletsen, je twee heerlijke kinderen en natuurlijk Erik, ik geniet elke keer weer, ik hou van jullie! Je ziet snel die witte muis wel weer het erf opdraaien. Lieve Han, waar zal ik beginnen? Het Monsterse strand, zeilkampen of Uganda. Wat geniet ik van onze vriendschap, de gesprekken, het reizen en de humor. I Love you! Janssie, jaar 1 geneeskunde, Heemskerkstraat: Janske en Babs aan het leren. Naast het doornemen van de fysiologie van het hart hadden we het ook vaak over wat er in ons eigen hart omgaat. En nog steeds, hoe waardevol. En wat zijn wij aan elkaar gewaagd als het gaat om afzien! Marmotte 2014? Renate en Michelle, 'sisters'. Thanks for the prayertime and friendship. Jasper, BTM roomie, al woon je nu wat verder met je heerlijke gezin, je blijft nog steeds mijn favoriete neef!

Len, Floor en Jet, mijn lieve zussen. Wat hou ik veel van jullie en wat ben ik trots op jullie. Floor staat hier als paranimf, maar eigenlijk staan jullie hier alle drie! Ik geniet van de momenten dat we bij elkaar zijn. Jullie kennen mij als geen ander. His en Willem, jullie horen daar uiteraard ook bij. Juda, nooit geweten dat je zoveel van een kleine jongen kan houden! Wat gaan we nog veel van jou genieten.

Lieve pap en mam, wat hou ik ontzettend veel van jullie. Een 'thuis' hebben is heel waardevol. Bedankt voor alle liefde, bemoediging en gezelligheid.

All Glory to Him

Addendum

Differential susceptibility to early literacy intervention in children with mild perinatal adversities: short and long-term effects of a randomized control trial

V.A.C. van der Kooy-Hofland, J. van der Kooy, A.G. Bus, M.H. van IJzendoorn, G.J. Bonsel

Journal of Educational Psychology, 2012:p337-349

ABSTRACT

In a randomized control trial we test whether short- and long-term effects of an early literacy intervention are moderated by mild perinatal adversities in accordance with differential susceptibility theory. One-hundred five-year-old children (58 percent male), who scored at or below the 30th percentile on early literacy measures were randomized to a web-based remedial early literacy program *Living Letters* or a treated control group. Parents gave written informed consent to access the perinatal data of their children at the Perinatal Register in the Netherlands. Twenty-one children were at birth small for gestational age (between the 2.5th - 10th percentiles) or late preterm (between 34 - 37 weeks + 6 days). In this group with mild perinatal adversities, intervention children outperformed the control group immediately after the intervention and after eight months of formal reading instruction, but a similar effect of the computerized literacy program in children without mild perinatal adversities was absent. In line with the theory of differential susceptibility children with mild perinatal adversities seem to be more open to environmental input, for better *and* for worse.

INTRODUCTION

Mild perinatal adversities such as being small for gestational age or being born late preterm are usually considered to be risk factors for subsequent child development, including cognitive development.[1-3] Here we present experimental data supporting a radically different view, derived from the theory of differential susceptibility.[4-6] We suggest that mild (but not severe) perinatal adversities may have programmed children to be more susceptible than other children to the environment, for better *and* for worse.

Children who are small for gestational age or born late preterm may acquire poorest early literacy skills in unfavorable environments but they might perform at the highest literacy level if delays in early literacy development are addressed at an early stage. Children may profit more from beginning reading instruction when they have received an early literacy intervention in kindergarten that prompted them to pay attention to print as an object of exploration – an important precursor of the beneficial effects of reading instruction.[7-11]

For kindergarten children with early literacy-related delays we tested in the current randomized control trial (RCT) whether after the intervention and at the end of first grade children with mild perinatal adversities were differentially susceptible to an early computer-based literacy intervention.

The impaired neuromotor, medical, social and cognitive development of very preterm children (< 32 weeks) and children who are extremely small for gestational age (< 2.5th percentile) has been extensively documented.[12-15]

For children with mild perinatal adversities fewer studies on their later development have been conducted. However, most available evidence supports the conventional assumption that children with mild perinatal adversities are at risk for medical and neurocognitive problems as well. Compared to full term children, children born late preterm had lower reading scores[1,16,17], a two times higher risk for special education at all grade levels[3], and a greater risk for developmental delays and school-related problems.[18] Cognitive development of mild small for gestational age children has also been shown to lag behind, with increased risk of learning disabilities and impaired learning-related abilities in childhood, and lower educational achievement among adults born near-term.[2,16,19] No evidence is available however on the effectiveness of enriched educational environments created for children with previous mild perinatal adversities.

In developmental psychopathology the concept of 'biological sensitivity to context' or more general 'differential susceptibility' has emerged to acknowledge the accumulating evidence that some children with a specific neurobiological, temperamental or genetic make-up

seem to suffer most from negative environments but at the same time appear to profit most from positive environments, for better *and* for worse.[5,6,20-22] The core idea is that not every child seems equally susceptible to the same parental, educational or environmental influences. Temperament has been one of the differential susceptibility factors central in the first wave of studies pioneered by Belsky and colleagues.[23] For example, an intervention that provided both high quality child care and parenting support showed a moderating effect of infant negative emotionality with respect to subsequent cognitive functioning and externalizing behavior.[24] A reactive temperament seems not a 'risk' but a susceptibility factor. Genetic differential susceptibility has been introduced by a Leiden group[25] who documented the potential role of dopamine-system genes for differential susceptibility. For example, children with the DRD4 7-repeat allele and unresponsive mothers displayed more externalizing behavior problems than children without the DRD4 7-repeat variant (irrespective of maternal responsiveness); but children with the DRD4 7-repeat allele and responsive mothers showed the lowest levels of externalizing problem behavior.[26]

Physiological factors (i.e., biological reactivity) have been introduced by Boyce and his team.[27] In a pioneering study on biological sensitivity to context Boyce et al[27] showed that 3-5 year old children with low cardiovascular or immune reactivity to stressors had approximately equal rates of respiratory illnesses in both low and high adversity settings. Highly biologically reactive children exposed to high adversity child care settings or home environments had substantially higher illness incidences than all other groups of children. Unexpectedly, they also found that highly sensitive children living in more supportive child care or family settings had the *lowest* illness rates, lower than even low reactivity children in comparable settings (see Ellis et al[6] for an extensive review of converging evidence). Here we suggest that mild perinatal adversities may have been associated with physiological changes such as higher cardiovascular reactivity to context, which according to the study of Boyce and colleagues would make children more sensitive to context, for better *and* for worse. Because of their stress reactivity, children with mild perinatal adversities may easily shut themselves off from learning experiences in a less optimal learning environment, whereas they might be most eager to learn from positive feedback in a supportive learning environment.

We present the first educational intervention study using a randomized control trial to demonstrate the short- and long-term, high learning potential of children with mild perinatal adversities in an optimal educational environment. The study targets a literacy intervention developed as a remedial program for children who lag behind in early literacy skills and who therefore are at risk not to benefit optimally from beginning reading instruction.[28,29] *Living Letters*, a computer-based educational program, compensates for a lack of environmental experience that promotes early literacy skills and once children have acquired these

competencies that are fundamental for learning to read they are better able to benefit from formal reading instruction in second grade.[10] In line with the theory that mild perinatal adversities are not a 'risk' but a susceptibility factor we expect that children with mild perinatal adversities will outperform the children without adversities when they receive the program but lag further behind without program. Due to an early, preventive remedial program at kindergarten, children with mild perinatal adversities may be better prepared for further reading instruction, and because of the hierarchical nature of the reading process early interventions may reveal long-term effects on their school achievements as well.[30]

The intervention program in this study uses the proper name to provide surface perceptual features of letters that help children discover sound-symbol relations between the first letter of their name and its sound in its spoken counterpart.[10,11] There is compelling evidence showing that name writing is commonplace in young children's everyday life and that the proper name is one of the first perceptually familiar words to young children.[31,32] By calling children's attention to sounds of letter units in the written name (e.g., "It's /pi/ of Peter") children receive a substantial amount of direct instruction about letters as symbols for sounds in the name. Most kindergarten children begin to combine understanding of how a word sounds with knowledge of how a word looks by using opportunities for development enhancement in daily life.[33,34] An individualized remedial computer program, *Living Letters,* was modeled after literate home activities with the proper name as a crucial prompt to stimulate combining understanding of what the name looks like with knowledge of how the name sounds. The computer program is especially created for children lacking in competencies fundamental to reading success who easily shut themselves off for learning experiences at home and in school. We expect the program to be more successful in holding these children's attention by providing constructive feedback immediately following an error [35] as well as by being adaptive to characteristics of the user or to the user's interaction with the system.[36] For instance, the program offers more feedback (more cues for solving the task) when a child fails the task and help is reduced when the learner is more competent and solves problems after a few attempts.

Aims and hypotheses

In the current randomized control trial we include 100 five-year-olds who scored at the lowest level of early literacy skills in the fall of the senior kindergarten year. Our central question is whether mild perinatal adversities moderate effects of a remedial intervention program targeting kindergarten children lacking in competencies fundamental to their school success – notably in the area of literacy. Findings so far show that the overall effects of Living Letters are moderate immediately after the intervention as well as over the

long-term.[10,11] Such a pattern of findings may manifest itself because the intervention increases learning only for those children who are most susceptible to their environment and need systematic instruction and support to explore print.

Children with mild perinatal adversities may be more susceptible to the environment including compensatory educational intervention in kindergarten and outperform children with mild perinatal adversities who do not receive the compensatory educational intervention as well as children without mild perinatal adversities who received the intervention. This differential intervention effect would emerge not only directly after the intervention but also a year later. Due to a better starting position as a result of the intervention, children with mild perinatal adversities benefit more from beginning reading instruction in first grade. Without a timely enriched environment, literacy performance of these children is expected to remain at a lower level because they, in contrast to their peers without perinatal adversities, are less receptive to influences in their daily environment. In line with differential susceptibility theory we expect therefore that the majority of children without perinatal adversities are less susceptible to compensatory education in kindergarten. Over the long-run they remain at roughly the same level whether or not they received a remedial computer program in kindergarten, and they will be outperformed by their peers with mild perinatal adversities who participated in the enriched literacy environment. The predicted differential effectiveness of Living Letters should be independent of possible contaminating systematic SES, IQ or executive function differences between the groups.

The current study aims therefore at testing the following hypotheses: (1) Children with mild perinatal adversities are most susceptible to early compensatory interventions in kindergarten and show the strongest beneficial effects directly after the intervention. (2) Beneficial effects stretch to the period of beginning reading instruction that builds on early literacy skills. Short-term positive effects assessed directly after the intervention in the mild perinatal adversities group are predicted to be maintained for reading tests at the end of grade 1 almost one year after the end of the intervention.

METHODS

Participants

The intervention sample was drawn from 15 regular public schools with a "normal" population. Eligible for participation were pupils speaking Dutch as their first language and between 60 to 72 months old. An estimated 12% of 452 pupils did not participate in

the screening, due to illness or absence for other reasons or failure of parental consent. The lowest scoring 30% ($N = 135$) on an aggregate measure composed of three screening assessments (letter knowledge, writing 'mama' and writing other words) was selected to participate in the experiment. The cut-point of 30% was based on the finding that this selection encompassed all children who knew very few letters and were unable to represent letters phonetically in their writings. The selected children were randomly assigned (ratio 2:1) to the intervention program (Living Letters) or treated control group (Living Books) stratified for school and gender.

One-hundred parents (74%) gave written informed consent to access the perinatal data of their children at the Perinatal Register in the Netherlands (PRN, 2010). The children (58 percent male) were at the start of the study 60 to 71 months old ($M = 64.16$ months, $SD = 2.99$). Almost all schools (14 out of 15) were represented in the subsample. Children were assigned to the group with mild perinatal adversities a priori defined as birth weight ranging between the 2.5th and 10th percentile for the gestational age (full term small for gestational age), or being late preterm that is a gestational age at birth of 34-37 weeks + 6 days. Twelve children (eight in intervention program, see Table 1) were at birth small for gestational age and nine late preterm (five in intervention program).

The children ($N = 100$) participating in the current, perinatal part of the intervention study did not significantly differ from the total sample ($N = 135$) on educational level of the father, verbal intelligence, regulatory skills, pretest early literacy skills, and outcome measure for early literacy skills. In grade 1 we lost seven children in the no mild adversities group (three in intervention program), because families moved ($n = 3$) or children duplicated senior kindergarten classroom ($n = 4$).

Study Design

A randomized pretest-posttest control group design was used to examine the differential effects of a remedial intervention (*Living Letters*) in kindergarten. Control subjects were assigned to another computer program not focusing on early literacy skills (*Living Books*). Eligible children were randomly assigned to intervention and control group, stratified for school and gender.

To examine whether randomization had been successful, we applied t-tests with experimental group (Living Letters and control group) as a factor and with mild adversities versus no mild adversities as a factor to test whether they were similar on paternal education, verbal and non-verbal intelligence, regulatory skills and pretest early literacy skills (Tables 1 and 4).

Table 1 Background variables as a function of experimental group and perinatal adversities

	Mild perinatal adversities (n = 21)		No mild perinatal adversities (n = 79)	
	LL (n = 13)	Control (n = 8)	LL (n = 52)	Control (n = 27)
Gender (male)	7	6	30	15
Paternal education	4.31 (2.29)	3.38 (1.85)	5.85 (2.19)	5.00 (2.04)
PPVT [a]	77.23 (8.90)	74.50 (13.59)	81.35 (12.13)	78.07 (11.28)
Raven's CPM [a]	15.15 (4.20)	18.00 (3.70)	16.92 (3.29)	16.85 (4.14)
Regulatory skills [b]	–.08 (.76)	–.76 (1.33)	.17 (.99)	–.07 (.95)

Note. [a]raw scores. [b]z-score.

There were no significant differences between the experimental groups on paternal education, verbal and non-verbal intelligence, regulatory skills, and pretest early literacy skills.

Children of fathers with lower educational level were overrepresented in the mild adversities group (M = 5.56, SD = 2.17), compared to the no mild adversities group (M = 3.95, SD = 2.13), t = 3.03, df = 98, p < .01 (Table 1). Also pretest early literacy skills showed a significant difference between the mild adversities group (M = -.45, SD = .75) and the no mild adversities group (M = .10, SD = 1.06), t = 2.23, df = 98, p <.05 (Table 4). There were no significant differences on verbal or non-verbal intelligence and regulatory skills.

Intervention program

Living Letters

Living Letters, designed by a team of computer experts, designers, and experts in the field of education, and available for schools and parents via subscription, is aimed at training basic literacy skills. The child's proper name or another familiar name such as 'mama' [mom] [37] is used to illuminate how letters in names relate to sounds.[38,39] Since the proper name is often the first word that young children can read and write, children received the program version with the proper name unless the name's spelling was inconsistent with Dutch orthography (e.g., Chris or Joey). In those cases, the program used 'mama', another often-known name, as the target word.[40] Of the 40 games, 22 games provided practice in recognizing the proper name. Six games focused on recognition of the first letter of the proper name, and another 12 games provided practice in identifying pictures that start or end with the first letter of the child's name.

The sessions started with an attractive animation to explain the upcoming games; for instance, the two main characters, *Sim* and *Sanne*, discuss their name and discover that these names start with the same sound. Errors when solving the games are followed by increasingly supportive computerized oral feedback. Unlike most computer games, the program *Living Letters* gives adult-like feedback that goes beyond "great" or "not quite right, try again". First, the task is repeated (*Find the word that starts with the same sound as your name*), next a clue is given (*Which word starts with /t/ of Tom?*), and lastly the correct solution is demonstrated (*You hear the first sound of your name, /t/ of tom, in tent)*. After a maximum of three trials per assignment, in both conditions, Sim, Sanne, and the teddy bear start dancing to mark the end of an assignment, whether or not the child has given the correct answer, after which the next game starts.

Living Books

The control group was given an alternative computer treatment however not targeting letter-sound knowledge: Living Books. This program consists of five age-appropriate computerized books that include oral narration and video representations of the scenes, but no printed text, thus allowing the child to *read by listening*. In each 10-minute session, children read one electronic storybook and responded to four follow-up questions among which two about difficult words (e.g., *What are paving stones?*) and two about story events (e.g., *Is dad happy or angry?*) by choosing one out of three pictures. Each book was repeated three times across the 15 sessions. In each repeated reading, children responded to four new questions, totaling 12 questions per book.

Training procedure

Training sessions were held over a period of 15 weeks. Children spent an estimated 10 minutes per session playing Living Letters or Living Books. Sessions occurred during the morning either in the classroom or the computer room conditional upon the school routines. Children wore headphones to reduce noise and distractibility. Because the intervention was the first tryout of Living Letters, university students at the master's level were present to prevent or solve technical problems with the help of an off-site helpdesk. It was their task to log children in on the website and provide supervision and assistance to ensure that children could complete all sessions. However, they did not provide guidance in explaining or solving the computer assignments. The system stored which assignments each child had completed and the correct game automatically appeared on screen when the supervisor entered the child's name. The system was also programmed in a way that the session automatically discontinued after four games so that sessions had the same duration

and the program was held over a similar period. Each child thus played all games as often and in the same order. The system registered which assignments children had completed which enabled the main researcher to note a missed session and repeat that within one week. Thanks to the computerized treatment, fidelity checks were maximal.

Measures

Perinatal variables

Data from the Netherlands Perinatal Registry (PRN) 2000–2001 were used. The PRN is a database that contains the linked data from three registries: the national obstetric database by midwives, the national obstetric database by gynaecologists, and the national neonatal/pediatric database.[41] The PRN registry contains comprehensive data on pregnancy, provided pregnancy care (interventions, referrals), and pregnancy outcomes. The coverage of the PRN is about 96% of all deliveries in the Netherlands. The health care provider recorded all variables during prenatal care, delivery and neonatal and lying-in period. The data are annually sent to the national registry office, where a number of range and consistency checks are conducted. Criteria for assignment to the group with mild perinatal adversities were birth weight between the 2.5th - 10th percentile for the gestational age (small for gestational age) or late preterm that is a gestational age at birth of 34 - 37 weeks + 6 days.

Parental education was surveyed using the following scale of highest form of education completed by the fathers and mothers: 1 (primary school), 2 (preparatory secondary vocational education), 3 (preparatory middle-level vocational education), 4 (senior secondary vocational education), 5 (senior secondary education), 6 (pre-university education), 7 (professional higher education), and 8 (university). Because the measures were strongly correlated but paternal education was more strongly associated with perinatal adversities we preferred this measure to maternal education as a covariate.

Intelligence. To control for intelligence as a confounding factor we tested verbal and non-verbal intelligence with the Dutch version of the Peabody Picture Vocabulary Test and the Dutch version of Raven's Colored Progressive Matrices.[42]

Regulatory Skills. Because regulatory skills relate to learning via the computer [43] we assessed regulatory skills at pretest with four tasks: (1) Following the Stroop paradigm, children had to switch rules by responding with an opposite, i.e., saying "blue" to a red dog and "red" to a blue dog.[44] The task consisted of 96 trials distributed over four conditions, in which demands on working memory (remembering the name of one or two dogs) and inhibition of the most obvious response (e.g., saying "blue" to a red dog) varied. Incorrect naming and

corrections were both scored as errors. (2) In a second Stroop-like task (opposites) children had to respond with the opposite to contrasting pairs of pictures [e.g., saying "fat" to thin] (based on Berlin at al.[45]). Incorrect naming and corrections were both scored as errors. (3) In a test called 'same tapping' the child copied the experimenter's hammer taps on cubes.[46] Each correct imitation in this working memory task was awarded one point with a maximum score of 12. (4) In the peg tapping test the child tapped twice with a pencil after one tap by the experimenter, and vice versa.[47] The total score was the number of correct responses to 16 items. Intraclass correlation coefficients between two independent coders were high for all four tasks ($r > .97$). PCA revealed one component with high loadings (.66 - .75) explaining 49% of the variance. The distribution of this aggregated measure (regulatory skills) was normal for both the intervention and the control group.

Screening tests

Screening tests aimed at identifying kindergarten children delayed in the basic understanding that letters relate to sounds. Rhyming did not discriminate in this age-group, however the following measures did:

Early writing. Children were asked to write familiar words like *mama* (mom) and four other words (e.g., *boot* [boat]).[46] Each word was double-coded on a scale from 1 (writing-like scribbles) to 6 (conventional spelling).[32] A score of 3 or higher indicates that one or more letters are represented phonetically. The intra-class correlation coefficient for 20 double-coded assignments was high ($r = .99$).

Receptive letter knowledge task. Children were asked to point to one of eight target letters, each presented on a card between four other letters.

Aggregated screening score. Alpha reliabilities for the tests were satisfactory; see Table 2. Correlations among the tests were rather high (> .52). Principal Component Analysis (PCA) on the measures revealed one component with high loadings (.83 - .87) that explained 74% of the variance. The lowest scoring 30% on the composite measure were selected as experimental group because they did not represent letters phonetically.

Early Literacy Skills

To test whether the program stimulates and re-organizes attention to sounds and letters in spoken and printed words the following test battery was applied (NELP, 2008):

Phonological skills (pre- and post-tested) were assessed in a series of 5 tasks: (1) identifying among three words the one that starts with a sound different from the other two words;

(2) selecting among four words two words with the same initial sound; (3) selecting from four words two words with the same final sound; (4) naming the first sound of words; and (5) naming all sounds of words. To reduce examiner bias, all picture names were pronounced by a computerized voice. All target sounds (n = 20) were consonants; all words were monosyllabic (CVC or CVVC). Each correct response was awarded one point (maximum = 25).

Word Recognition (only post-tested). Children had to identify the depicted target word (e.g., raam) among four printed words. The (incorrect) alternatives differed in 1 (room), 2 (rat) or all letters (bon) from the target word. Correct responses were rewarded with 3 points (raam); correspondence of first and last letter (room) with 2 points; correspondence of first letter with 1 point (rat); and no correspondence (bon) with 0.

Decoding (only post-tested). Children were trained in decoding four vowel-consonant (VC) and four consonant-vowel-consonant (CVC) nonsense words. If children failed to pronounce the nonsense word in the first five seconds after presentation of a word, they were stimulated to sound out the separate letters. If this did not prompt correct decoding, the experimenter pronounced the separate sounds and stimulated the subjects to blend the sounds. If they did not succeed, the experimenter repeated the separate sounds, blended them, and had subjects repeat the naming and blending. The list of eight words was repeated five times in different sequences. Scores per word varied from 5 (successful first attempt) to 1 (non-completion of item).

Early Literacy Skills. Alpha reliabilities for all tests were satisfactory; see Table 2. As is shown in Table 3 the three measures were strongly correlated (> .59). Principal Component Analysis (PCA) on the posttest measures revealed one component with high loadings (.81-.77) that explained 77% of the variance. This component was labeled as 'Early Literacy Skills' and used as dependent variable. The distribution of this aggregated measure was normal for both the intervention and the control group. We used phonological skills measured at pretest as covariate (see Table 4).

Beginning Reading Skills

In first grade, children were tested with measures that are used in Dutch schools to assess reading development after about 8 months of instruction. In a regular orthography such as the Dutch language, measures that target speed discriminate better than measures of accuracy.[48]

Table 2 Reliabilities and Mean scores (standard deviations) for screening-tests, child characteristics, pretests, and posttests

	Measure		*N*	*M*(*SD*)	Skewness(*SE*)	Kurtosis(*SE*)
Child characteristics	Paternal education		100	5.22 (2 25)	–.04 (.24)	–1.47 (.48)
	Peabody Picture Vocabulary Test	.94[b]	100	79.38 (11.70)	–.02 (.24)	–.44 (.48)
	Non-verbal intelligence	.69[c]	100	16.76 (3.70)	–.03 (.24)	–.06 (.48)
	Executive Functioning (factor score)		100	.00 (1.00)	–.46 (.24)	.12 (.48)
	- Stroop task dogs (max = 96)	.94[a]	100	83.98 (8.60)	–.89 (.24)	– 17 (.48)
	- Stroop task opposites (max = 48)	.91[a]	100	31.16 (7.44)	–.50 (.24)	–.18 (.48)
	- Peg Tapping (max = 16)		100	13.35 (2.32)	–.66 (.24)	–.41 (.48)
	- Same Tapping (max = 12)	.80[c]	100	6.88 (2.40)	–.10 (.24)	–.45 (.48)
Screening	Screening (aggregated score)		404	.00 (.86)	–.23(12)	.04 (.24)
	- Letter knowledge (0-8)	.73[a]	404	5.62 (2.07)	–.46(.12)	–.92 (.24)
	-Writing 'Mom' (1-6)		404	3.67 (1.74)*	.18 (.12)	–1.27 (.24)
	- Writing words (1-6)	.87[a]	404	3.04 (1.07)*	.23 (.12)	.82 (.24)
Pretest	Phonological skills (max = 25)	.81[a]	100	7.54 (4.86)	1.02 (.24)	.19 (.48)
Posttest	Early literacy skills (factor score)		100	.00 (1.00)	.44 (.24)	–.73 (.48)
	- Phonological skills (max = 25)	.87[a]	100	12.41 (6.03)	.04 (.24)	–1.20 (.48)
	- Word recognition (max = 45)	.80[a]	100	27.29 (6.55)	.32 (.24)	–.10 (.48)
	- Decoding (max = 40)	.98[a]	100	15.08 (4.53)	–.54 (.24)	–.68 (.48)
Post-posttest	Reading Grade 1 (factor score)		93	.00 (1.00)	14 (.25)	–.56 (.50)
	- Word reading fluency	.91[b]	93	22.52 (10.33)	28 (.25)	– 48 (.50)
	- Pseudo-word reading test	.88[b]	93	21.82 (11.94)	.71 (.25)	–.39 (.50)
	- Serial naming	.74[a]	93	40.73 (8.90)	.78 (.25)	–.05 (.50)

[a]Cronbach's alpha; [b]test-retest reliability; [c]split-half reliability; *scores ≥ 3 indicates writing one or more phonetic symbols.

Table 3 Bivariate correlations among measures

	Time	1[a]	2[a]	3[a]	4[a]	5[b]	6[b]	7[b]
1. Phonological skills	pretest	–	.55**	.51**	.65**	.25*	.18	–.25*
2. Phonological skills	posttest		–	.54**	.79**	.32**	.23*	–.29**
3. Word recognition	posttest			–	.64**	.31**	.30**	–.16
4. Decoding	posttest				–	.28**	.22*	–.21*
5. Word reading fluency	post-posttest					–	.93**	–.64**
6. Pseudo-word reading test	post-posttest						–	–.57**
7. Serial naming	post-posttest							–

** p < .01. *. p < .05. Note: [a]*n* = 100; [b]*n* = 93.

Table 4 Mean pretest and posttest scores (z-scores) and SE's as a function of experimental group and perinatal adversities

	Mild perinatal adversities (*n* = 21)		*P*	No mild perinatal (*n* = 79)[a]		*P*
	LL	Control		LL	Control	
	n = 13	*n* = 8		*n* = 52[b]	*n* = 27[c]	
Pretest						
Phonological skills	–.39 (.22)	– .55 (.24)	*ns*	.14 (.15)	.03 (.20)	*ns*
Posttest						
Early literacy skills (factor score)	.64 (.25)	–.47 (.11)	<.05	–.06 (.10)	–.06 (.16)	*ns*
Post-posttest						
Reading Grade 1 (factor score)	.42 (.21)	–.41 (.22)	<.05	–.04 (.11)	.00 (.24)	*ns*

Note: [a]Grade 1, *n* = 72; [b]Grade 1, *n* = 49; [c]Grade 1, *n* = 23.

Word reading fluency was tested with the one-minute-test, a standardized test, to determine how many words from a list can be read during one minute.[49]

Pseudo-word reading fluency, a standardized pseudo-word reading test, assessed how many nonsense words were read accurately in 2 minutes.[50]

Serial naming of letters. To assess how fast letters can be retrieved from memory children named 50 lowercase letters composed of five different letters (d, o, a, s and p) non-consecutively ordered as fast as possible.[51]

Aggregated measure for grade 1 reading. Cronbach's and Guttman's alpha's for the standardized tests were satisfactory; see Table 2. The measures for word reading, non-word reading and serial naming of letters showed high correlations (> .55). Principal Component Analysis revealed one component explaining 84% of variance with test loadings ranging from .87 to .95. The distribution of this aggregated variable was normal for both the intervention and the control group.

Data collection and scoring procedure

In fall (screening), one month before the 15-week intervention, directly after the intervention, and after 8 months of instruction master's level university students, blind to treatment, tested the children. Assessments were delivered in a fixed order to all participants. Examiners were extensively trained in administration procedures. Videotaped pre/post assessments were used to control the testing procedure. Master's level university students, blind to treatment, scored tests under supervision of the main researcher.

Analysis

Because the subjects were recruited from 14 schools and observations within schools may be dependent we started with deriving the Huber-White estimates to correct for clustering of the measures.[52] We then included these estimates in the Complex Sample General Linear Model (CSGLM, SPSS 17) to carry out regression analyses with reading skills directly after the program in kindergarten and at the end of first grade as dependent or outcome variables, and pretest early literacy skills, paternal educational level, children's PPVT score (verbal intelligence), Raven score (non-verbal intelligence), regulatory skills, presence of mild perinatal adversities, and experimental group (Living Letters versus control group) as covariates (total $N = 100$ children in 14 schools).

RESULTS

Paternal educational level, verbal and non-verbal intelligence and regulatory skills were included as covariates. None of these variables were significant predictors of beginning literacy skills in kindergarten and grade 1 reading skills. Experimental group and mild perinatal adversities group did not show significant main effects on early and beginning literacy skills. The interaction between experimental group and mild perinatal adversities, however, was significant not only immediately after the intervention with early literacy

skills as the dependent measure, F (1, 13) = 8.25, p = .013, but also at the end of grade 1 with reading kills as the dependent variable, F (1, 13) = 9.22, p < .01.[1&2]

In order to examine the interactions between intervention and mild perinatal adversities, we repeated the CSGLM in the mild perinatal adversities and the no mild perinatal adversities groups separately. The significant effect of experimental group in the mild perinatal adversities group for posttest early literacy skills (F (1, 8) = 7.24, p < .05; n = 21) was still present at the end of grade 1 reading (F (1, 8) = 5.79, p < .05; n = 21), where children in the *Living Letters* group outperformed the control group (Table 4). However, the no mild perinatal adversities group was not susceptible for the early intervention as was demonstrated by the absence of a significant effect for end of kindergarten early literacy skills (F (1, 13) = .06, p = .82; n = 79) and grade 1 reading skills (F (1, 13) = .01, p = .91; n = 72). Outcomes were basically the same when the small for gestational age and late preterm group were analyzed separately but for these post-hoc analyses statistical power was of course low due to the small number of subjects in the sub-groups.

In Table 5 the effect sizes of the intervention for early literacy skills and grade 1 reading are presented. Although the overall effect of the computer intervention was low, the children with mild perinatal adversities benefited substantially from the intervention. The effect size for reading in first grade (d = 1.11) was only slightly lower than the effect size for post test early literacy of d = 1.24 directly after the intervention. Figure 1 illustrates the difference in effect size between experimental groups for post test early literacy and grade 1 reading scores. The dependent measures are the aggregate scores on early literacy and grade 1

---

[1] A repeated measures approach was not accommodated by the Complex Sample General Linear Model in SPSS. Addressing the longitudinal design but ignoring the dependency of pupils' scores within schools, we ran a MANOVA with perinatal adversities and experimental group as factors and five covariates. This approach revealed a significant interaction between perinatal adversities and experimental group (F (1, 85) = 12.28, p < .001, η^2 = .126). The three-way interaction between point of measurement (directly after the intervention in kindergarten versus at the end of grade 1), perinatal adversities and experimental group was not significant (F (1, 85) = 0.05, p = .82). This indicates that the interaction between perinatal adversities and experimental group applied to both points of measurement.

[2] Because of high correlations between the separate tests, we used composite measures in the analysis. We ran the analysis for each of the six tests separately, aiming at excluding that only a few measures revealed significant interactions between intervention and perinatal adversities. The interaction was significant for 4 out of 6 tests and marginally significant for decoding in kindergarten and the pseudo-word reading test in first grade.

Table 5 Effects of treatment in the total group and in the subsample (mild perinatal adversities) directly after the intervention (posttest early literacy skills) and after one year reading instruction controlling for background (paternal education, PPVT, RCPM, regulatory skills and pretest)

Measure	*n*	Estimate (*SE*)	95%*CI B*	*t*	*p*-value	Cohen's *d*[d]
Total group						
Early literacy skills	100	.12 (.07)	–.04,.28	1.66[a]	.12	.34
Reading Grade 1	93	.08 (.13)	–.19,.36	.65[b]	.53	.14
Subsample with mild perinatal adversities						
Early literacy skills	21	.36 (.13)	.05,.67	2.69[c]	< .05	1.24
Reading Grade 1	21	.40 (.17)	.02,.78	2.41[c]	< .05	1.11

Note: [a]*n* = 100, *df* = 13; [b]*n* = 93, *df* = 13; [c]*n* = 21, *df* = 8; [d]For calculating Cohen's *d* we applied Thalheimer & Cook's (2002) formula 2*t*/√*n*-2.

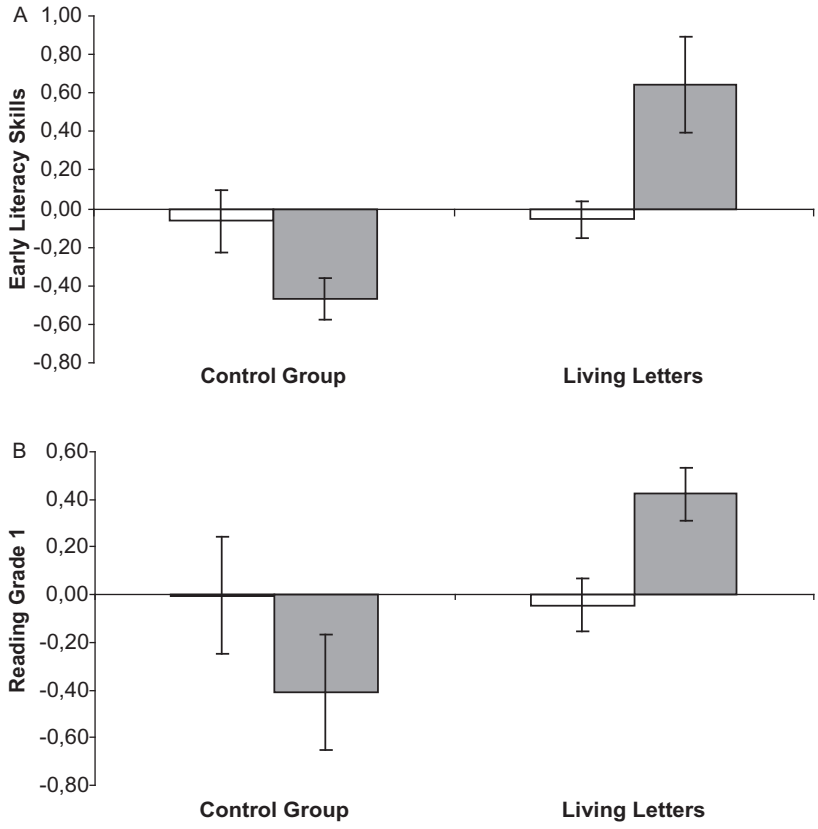


Figure 1 Estimated means and standard errors (error bars) for early literacy skills of children with mild perinatal adversities (grey) and without (white) in the intervention group (Living Letters) and in the control condition (A) directly after the intervention in kindergarten (N = 100) and (B) one year after the intervention in grade 1 (N = 93). *Note*. Grey: mild perinatal adversities. White: no perinatal adversities.

reading, residualized with the four covariates (paternal educational level, verbal intelligence, non-verbal intelligence, and regulatory skills) before computing means and standard deviations per sub-group. From Figure 1, it can be derived that the mild perinatal adversities group manifested the highest score on early literacy skills directly after the program and on grade 1 reading after the *Living Letters* intervention, and the lowest in the control group.

DISCUSSION

In this randomized control trial we found that, without an adequate preventive intervention program, children who had experienced mild perinatal adversities performed at the lowest level at the end of grade 1. In line with differential susceptibility theory, however, we also found that children with mild perinatal adversities profited most from a computer-based remedial intervention with an adaptive feedback regime, and these susceptible children kept their advantage even at the end of grade 1, after one year of formal reading instruction without any further additional support. As children in intervention and control condition were taught by the same teachers and exposed to similar classroom curricula we can be certain that there were no differences apart from the computer-based literacy intervention in kindergarten to explain this finding. In particular the advantage at the end of grade 1 demonstrates how important it is to address early literacy delays at an early stage.

Remediation of early literacy skills at an early stage can enhance effects of systematic instruction in reading skills that in the Netherlands does not begin until children are in first grade. We also found that some children are more susceptible to early remedial interventions than other children. Children with mild perinatal adversities are vulnerable to develop persistent delays in literacy skills but they also seem to thrive and are quick in acquiring high levels of elementary literacy skills when they have a chance to catch up and outrun their peers prior to the start of beginning reading instruction by participating in an enriched, computer-based literacy environment in kindergarten. These susceptible children seem to have not only risk factors but also unexpected learning potentials when a rearing environment includes elements that make children attentive to the basic ingredients of reading. For the children who did not suffer from perinatal adversities the intervention did not result in short- or long-term elevated levels of literacy skills.

In the dominant paradigm of developmental psychopathology the cumulative nature of risk factors has been emphasized, and the diverging developmental pathways of children with specific vulnerabilities in challenging environments.[53] Most of the developmental studies of the past few decades have focused on children at risk for deviant development because of a

combination of child-related and environmental risk factors. The prevailing tunnel view on risks prevented developmental researchers from paying equal attention to the other side of the coin, optimal development in supportive environments. Differential susceptibility theory draws attention to the possibility that in a wider view on environmental risks *and* positive contexts child-related risk factors might turn out to create greater susceptibility to positive environments. Reactive or difficult temperament has been one of the differential susceptibility factors central in the first wave of studies pioneered by Belsky et al[23] The potential role of dopamine-system related genes for differential susceptibility has been introduced by Bakermans-Kranenburg and Van IJzendoorn[25] for social-emotional and by Kegel et al[54] for cognitive development. Physiological factors (i.e., biological reactivity) have been introduced by Boyce and his team.[27]

From this latter, pediatric perspective Boyce and Ellis[21] used the metaphor of 'dandelions' to indicate the large number of children who are genetically or prenatally programmed in a way that they would survive and function rather robustly within almost any environment. The smaller number of 'orchids' however would wilt quickly in neglecting circumstances but bloom in a spectacular way when raised in the most optimal environment. What seems to be a risk factor in an average or bad environment, e.g. biological reactivity to stress or a reactive temperament, turns out to promote optimal development in a nurturing context. The 'orchid hypothesis'[55] misleadingly suggests two distinctive classes of individuals instead of a continuum of more or less susceptibility to the environment. The orchid metaphor however also brings home effectively the surprising message that seemingly maladaptive but evolutionary enduring traits might have an important adaptive role in specific niches as they contribute to the indispensable variation in the human species. This is the evolutionary view on differential susceptibility that Belsky[4,56] was the first to articulate.

Why would mild perinatal adversities be susceptibility factors instead of only risk factors? In general, susceptibility to context is associated with characteristics that enhance the individual's ability to monitor the environment and to extract most effectively its reward value. In Suomi's[57] studies on rhesus monkeys anxious, timid offspring would become anxious adults when they were reared in a harsh parenting environment but they would flourish and climb the hierarchy of the troop when they were raised by sensitive mothers who allowed them to use their anxious monitoring of the environment to elevate their access to resources. In a study of temperament and maternal discipline in relation to externalizing problems in early childhood, Van Zeijl et al[58] found that children with reactive temperaments were more susceptible to both negative and positive discipline than children of relatively easy temperament. On basis of their longitudinal studies Kochanska et al[59] proposed that

the ease with which stress and anxiety is induced in reactive children helps them to respond most favorably to gentle parental discipline but at the same time to be most vulnerable to the negative effects of harsh parenting.

Mild perinatal adversities may be associated with elevated levels of stress in the expectant mothers.[22] Although the number of studies on the association between maternal stress as indexed by HPA-axis functioning and mild perinatal adversities is relatively small and findings are equivocal, a recent large prospective cohort study suggested that larger cortisol awakening responses (CAR) in early pregnancy may be related to lower birth weight and higher small for gestational age risk.[60,61] Mild perinatal adversities may lead to higher cardiovascular and HPA-axis reactivity to context, which according to the pioneering study of Boyce and colleagues[27] would make children more sensitive to context, for better and for worse.

Low birth weight babies showed increased cortisol concentrations in umbilical cord blood and raised urinary cortisol excretion in childhood.[62] In adult life they have higher pulse rates, an index of sympathetic activity, and increased fasting cortisol concentrations.[63] Studies showed an enhanced plasma cortisol response to synthetic adrenocorticotrophic hormone.[64,65] An increased stress response was observed in low birth weight children.[66,67] Because of their elevated stress reactivity children with mild perinatal adversities may easily shut themselves off for learning experiences in a less organized and rewarding environment, whereas they might be most eager to learn from exposure to relevant experiences and positive feedback in a supportive learning environment.[22]

The *Living Letters* interactive computer program has been designed to reflect the interactions of sensitive parents guiding their preschoolers into the world of written language. Before formal reading instruction children already are curious to know how written words convey meaning, and how the visual form relates to the spoken counterpart of words. In particular their proper name is focus of this search that mostly starts at an early age far before reading instruction begins. Sensitive parents reinforce the discovery of alphabetic writing by encouraging interest in the proper name and other frequently used names and helping their children to combine their understanding of how the name looks with knowledge of how the name sounds. They are quick to recognize and reward successful attempts to name the first letter of the proper name and recognize the letter in other words.

The computer program *Living Letters* creates a similar type of sensitive and stimulating environment for the acquisition of early literacy skills. For example, as the proper name provides surface perceptual features that help children to discover relations between

letters of the name and sounds in spoken counterparts the program uses the child's own name to initiate the same discovery process. Errors when solving the games are followed by increasingly supportive computerized oral feedback of an adaptive and constructive nature. The feedback carefully scaffolds the children's trials and errors in their search for answers to the challenges of the games by repeating the task, giving clues, and finally demonstrating a correct solution. This is the optimal learning environment in which children with mild perinatal adversities seem to learn most effectively while children without perinatal adversities gain as much whether experiences with print are specific or nonspecific.

An important limitation of the current study is randomization to treatment and control group without stratification to mild perinatal adversities. Although we ascertained that the bias introduced by this lack of stratification was minimal as children with and without mild adversities were almost equally divided among treatment and control group, in future studies participants should also be randomized on the basis of the susceptibility factor –a limitation of all differential susceptibility experiments to date. Another limitation is the lack of information on the mechanism of susceptibility of children with mild perinatal adversities to the intervention. We speculated about the elevated stress reactivity of these children but HPA-axis or cardiovascular responses to stress were not assessed in the current study, and elevated stress reactivity remains a hypothesis to be tested in a future, more comprehensive trial. Lastly, we included only one potential marker of differential susceptibility, namely mild perinatal adversities in the current study although another study on a different sample but with the same *Living Letters* intervention program revealed a strong differential susceptibility effect for carriers of the 7 repeat dopamine D4 receptor gene (DRD4-7R, Kegel et al., 2011). To examine the associations between various susceptibility markers (temperament, regulatory skills, dopamine-related genes, perinatal adversities) in future studies these markers should be included simultaneously.

Our finding of enhanced susceptibility of children with mild perinatal adversities to the environment has at least two important implications. First, children with mild perinatal adversities have traditionally been viewed as being at risk for delays in later (cognitive) development. The current study shows that they may have a high potential for learning in enriched environments. What seems to be a risk factor turns out to be a potential asset which deserves our special attention in creating adequate educational environments. Second, experiments demonstrate that interventions might have only weak to modest effects on children's health or development across the board.[10,11,43,68] Effect sizes in this study for the main effect of the interaction disregarding the susceptibility factor, for instance,

remained below half a standard deviation. Yet the intervention appeared to be strongly effective for the specific, susceptible sub-group, as indexed by mild perinatal adversities. Even long-term effects of the intervention amounted to more than one standard deviation on an aggregate reading skills assessment that was in no way directly targeted in the computer-based early intervention program. For the majority the gap after 8 months of reading instruction was already present at enrollment in grade 1. In evaluating experimental interventions researchers should take differential effects of their manipulations into account as predicted by differential susceptibility theory.[6] The age-old intervention question of what works for whom might be fruitfully addressed from the perspective of differential susceptibility theory.

In sum, we found that children who experienced mild perinatal adversities might be at risk for early reading problems but in an enriched environment they may reach a high level of early reading skills, an advantage that still exists after one year of formal reading instruction. This provides unique experimental support for differential susceptibility theory in the cognitive domain and illustrates the double-edged nature of mild perinatal adversities as a risk factor for academic skills as well as a potential asset.

REFERENCES

1. Chyi LJ, Lee HC, Hintz SR, Gould JB, Sutcliffe TL. School outcomes of late preterm infants: special needs and challenges for infants born at 32 to 36 weeks gestation. J Pediatr. 2008 Jul;153(1):25-31.
2. Nomura Y, Halperin JM, Newcorn JH, et al. The risk for impaired learning-related abilities in childhood and educational attainment among adults born near-term. J Pediatr Psychol. 2009 May;34(4):406-18.
3. van Baar AL, Vermaas J, Knots E, de Kleine MJ, Soons P. Functioning at school age of moderately preterm children born at 32 to 36 weeks' gestational age. Pediatrics. 2009 Jul;124(1):251-7.
4. Belsky J. Differential susceptibility to rearing influence: An evolutionary hypothesis and some evidence. In B. Ellis & D. Bjorkland (Eds.), Origins of the social mind: Evolutionary psychology and child development (pp. 139-163). New York: Guilford. 2005.
5. Belsky J, Bakermans-Kranenburg MJ, Van Ijzendoorn MH. For better and for worse. Differential susceptibility to environmental influences. Current Directions in Psychological Science, 16, 300-304. doi: 10.1111/j.1467-8721.2007.00525.x. 2007.
6. Ellis BJ, Boyce WT, Belsky J, Bakermans-Kranenburg MJ, van Ijzendoorn MH. Differential susceptibility to the environment: an evolutionary—neurodevelopmental theory. Dev Psychopathol. 2011 Feb;23(1):7-28.
7. Byrne BM, Fielding-Barnsley R, Ashley L. Effects of preschool phoneme identity training after six years: Outcome level distinguished from rate of response. Journal of Educational Psychology, 92, 659-667. doi: 10.1037/0022-0663.92.4.659 2000.
8. Duursma E, Augustyn M, Zuckerman B. Reading aloud to children: the evidence. Arch Dis Child. 2008 Jul;93(7):554-7.
9. Silva C, Alves Martin M. Phonological skills and writing of pre-syllabic children. Reading Research Quarterly, 37, 466-483. doi: 10.1598/RRQ.37.4.6 Snider, V. E. (1997). The relationship between phonemic awareness and later reading achievement. Journal of Educational Research, 90, 203-211. Retrieved from http://www.jstor.org/stable/27542094 2002.

10. Van der Kooy-Hofland VAC, Bus AG, Roskos KA. Effects of a brief but intensive remedial computer intervention in a sub-sample of kindergartners with early literacy delays. Reading and Writing. Doi: 10.1007/s11145-011-9328-5. 2011.

11. Van der Kooy-Hofland VAC, Kegel CAT, Bus AG. Evidence-based computer interventions targeting phonological awareness for young students at-risk to prevent reading problems. In S. B. Neuman & D. K. Dickinson (Eds.), Handbook of Early Literacy. (Vol. 3, pp.214-227). New York: The Guildford Press. 2011.

12. Aarnoudse-Moens CS, Weisglas-Kuperus N, van Goudoever JB, Oosterlaan J. Meta-analysis of neurobehavioral outcomes in very preterm and/or very low birth weight children. Pediatrics. 2009 Aug;124(2):717-28.

13. Anderson P, Doyle LW, Victorian Infant Collaborative Study G. Neurobehavioral outcomes of school-age children born extremely low birth weight or very preterm in the 1990s. JAMA. 2003 Jun 25;289(24):3264-72.

14. Bhutta AT, Cleves MA, Casey PH, Cradock MM, Anand KJ. Cognitive and behavioral outcomes of school-aged children who were born preterm: a meta-analysis. JAMA. 2002 Aug 14;288(6):728-37.

15. Rodrigues MC, Mello RR, Fonseca SC. Learning difficulties in schoolchildren born with very low birth weight. J Pediatr (Rio J). 2006 Jan-Feb;82(1):6-14.

16. Kirkegaard I, Obel C, Hedegaard M, Henriksen TB. Gestational age and birth weight in relation to school performance of 10-year-old children: a follow-up study of children born after 32 completed weeks. Pediatrics. 2006 Oct;118(4):1600-6.

17. Lee ES, Yeatman JD, Luna B, Feldman HM. Specific language and reading skills in school-aged children and adolescents are associated with prematurity after controlling for IQ. Neuropsychologia. 2011 Apr;49(5):906-13.

18. Morse SB, Zheng H, Tang Y, Roth J. Early school-age outcomes of late preterm infants. Pediatrics. 2009 Apr;123(4):e622-9.

19. Johnson EO, Breslau N. Increased risk of learning disabilities in low birth weight boys at age 11 years. Biol Psychiatry. 2000 Mar 15;47(6):490-500.

20. Belsky J, Pluess M. Beyond diathesis stress: differential susceptibility to environmental influences. Psychol Bull. 2009 Nov;135(6):885-908.

21. Boyce WT, Ellis BJ. Biological sensitivity to context: I. An evolutionary-developmental theory of the origins and functions of stress reactivity. Dev Psychopathol. 2005 Spring;17(2):271-301.

22. Pluess M, Belsky J. Prenatal programming of postnatal plasticity? Dev Psychopathol. 2011 Feb;23(1):29-38.

23. Belsky J, Hsieh KH, Crnic K. Mothering, fathering, and infant negativity as antecedents of boys' externalizing problems and inhibition at age 3 years: differential susceptibility to rearing experience? Dev Psychopathol. 1998 Spring;10(2):301-19.

24. Blair C. Early intervention for low birth weight, preterm infants: the role of negative emotionality in the specification of effects. Dev Psychopathol. 2002 Spring;14(2):311-32.

25. Bakermans-Kranenburg MJ, van Ijzendoorn MH. Gene-environment interaction of the dopamine D4 receptor (DRD4) and observed maternal insensitivity predicting externalizing behavior in preschoolers. Dev Psychobiol. 2006 Jul;48(5):406-9.

26. Bakermans-Kranenburg MJ, van Ijzendoorn MH. Research Review: genetic vulnerability or differential susceptibility in child development: the case of attachment. J Child Psychol Psychiatry. 2007 Dec;48(12):1160-73.

27. Boyce WT, Chesney M, Alkon A, et al. Psychobiologic reactivity to stress and childhood respiratory illnesses: results of two prospective studies. Psychosomatic Medicine, 57, 411-422. Retrieved from http://www.psychosomaticmedicine.org. 1995.

28. Shonkoff JP, Phillips DA. From Neurons to Neighborhoods : The Science of Early Childhood Development. Washington, DC: National Academy Press. 2000.

29. Stipek DJ, Ryan RH. Economically disadvantaged preschoolers: ready to learn but further to go. Dev Psychol. 1997 Jul;33(4):711-23.

30. Heckman JJ. Skill formation and the economics of investing in disadvantaged children. Science. 2006 Jun 30;312(5782):1900-2.

31. Levin I, Both-de Vries AC, Aram D, Bus AG. Writing starts with own name writing: From scribbling to conventional spelling in Israeli and Dutch children. Applied Psycholinguistics, 26, 463-477. doi: 10.1017.S0142716405050253. 2005.

32. Levin I, Bus AG. How is emergent writing based on drawing? Analyses of children's products and their sorting by children and mothers. Developmental Psychology, 39, 891-905. doi: 10.1037/0012-1649.39.5.891. 2003.

33. Levin I, Aram D. Children's names contribute to early literacy: A linguistic and social perspective. In D. Ravid & H. Bat-Zeev Shyldkort (Eds.), Perspectives on language and language development (pp. 219-239). Dordrecht, the Netherlands: Kluwer. 2005.

34. Molfese VJ, Beswick J, Molnar A, Jacobi-Vessels J. Alphabetic skills in preschool: a preliminary study of letter naming and letter writing. Dev Neuropsychol. 2006;29(1):5-19.

35. Corbett AT, Anderson JR. Locus of feedback control in computerbased tutoring: Impact on learning rate, achievement and attitudes. In J. Jacko, A. Sears, M. Beaudouin-Lafon, & R. Jacob (Eds.), Proceedings of the ACM CHI'2001 Conference on Human Factors in Computing Systems (pp. 245–252). New York: ACM Press. . 2001.

36. Vasilyeva E. Towards personalized feedback in educational computer games for children: Review of recent research. Proceedings of IASTED WBE'2007 Conference, 597-602. Retrieved from http://www.iasted.org. 2007.

37. Levin I, Shatil-Carmon S, Asif-Rave O. Learning of letter names and sounds and their contribution to word recognition. Journal of Experimental Child Psychology, 93, 139–165. doi: 10.1016/j.jecp.2005.08.002. 2006.

38. Bus AG, Van Ijzendoorn MH. Phonological awareness and early reading: A meta-analysis of experimental training studies. Journal of Educational Psychology, 91, 403-414. Retrieved from http://www.apa.org/pubs/journals/edu. 1999.

39. Ehri LC, Nunes SR, Willows DM, Schuster B, Yaghoub-Zadeh Z, Shanahan T. Phonemic awareness instruction helps children learn to read: Evidence from the National Reading Panel's meta-analysis. Reading Research Quarterly, 36, 250–287. Retrieved from http://www.reading.org/general/publications/journals/rrq.aspx. 2001.

40. Both-de Vries AC, Bus AC.). Name writing: A first step to phonetic writing? Does the name have a special role in understanding the symbolic function of writing? Literacy Teaching and Learning, 12, 37–55. Retrieved from http://www.readingrecovery.org/rrcna/journals/ltl/index.asp. 2008.

41. Meray N, Reitsma JB, Ravelli AC, Bonsel GJ. Probabilistic record linkage is a valid and transparent tool to combine databases without a patient identification number. J Clin Epidemiol. 2007 Sep;60(9):883-91.

42. Van Bon W. Raven's Coloured Progressive Matrices. Nederlandse normen en enige andere uitkomsten van onderzoek [Raven's Coloured Progressive Matrices. Dutch norms and some other research findings]. Lisse, the Netherlands: Swets & Zeitlinger. 1986.

43. Kegel CAT, Van der Kooy-Hofland VAC, Bus AG. Improving early phoneme skills with a computer program: Differential effects of regulatory skills. Learning and Individual Differences, 19, 549-554. doi: 10.1016/j.lindif.2009.07.002. 2009.

44. Beveridge M, Jarrold C, Pettit E. An experimental approach to executive fingerprinting in young children. Infant and Child Development, 11, 107-123. doi: 10.1002/icd.300. 2002.

45. Berlin L, Bohlin G, Nyberg L, Janols LO. How well do measures of inhibition and other executive functions discriminate between children with ADHD and controls? Child Neuropsychol. 2004 Mar;10(1):1-13.

46. Schrooten JJF, Van Alphen de Veer RJ. Leidse Diagnostische Test, LDT. Amsterdam, the Netherlands: Swets & Zeitlinger.

47. Diamond A, Taylor C. Development of an aspect of executive control: development of the abilities to remember what I said and to "do as I say, not as I do". Dev Psychobiol. 1996 May;29(4):315-34.

48. Share DL. On the Anglocentricities of current reading research and practice: the perils of overreliance on an "outlier" orthography. Psychol Bull. 2008 Jul;134(4):584-615.

49. Brus BT, Voeten MJM. Een-Minuut-Test [One minute test]. Nijmegen, Netherlands: Berkhout Testmateriaal. 1973.

50. Van den Bos KP, Lutje Spelberg HC, Scheepstra AJM, De Vries JR. De Klepel. Een Test voor de Leesvaardigheid van Pseudo-woorden [The 'Klepel': A test for assessing the ability to read pseudo-words]. Nijmegen, the Netherlands: Berkhout Testmateriaal. 1994.

51. Van den Bos KP, Lutje Spelberg HC, Scheepstra AJM, De Vries JR. Snelheid van Benoemen en Woorden Lezen [Speed of serial naming and word reading]. Nijmegen, the Netherlands: Berkhout Testmateriaal. 2003.

52. Hatcher PJ, Hulme C, Miles JN, et al. Efficacy of small group reading intervention for beginning readers with reading-delay: a randomised controlled trial. J Child Psychol Psychiatry. 2006 Aug;47(8):820-7.

53. Cicchetti D. Developmental psychopathology: Reactions, reflections, projections. Developmental Review, 13, 471. doi: 10.1006/drev.1993.1021. 1993.

54. Kegel CAT, Bus AG, Van Ijzendoorn MH. Differential susceptibility in early literacy instruction through computer games: The role of the Dopamine D4 Receptor Gene (DRD4). Mind, Brain, and Education, 5, 71-78. doi: 10.1111/j.1751-228X.2011. 01112.x. 2011.

55. Dobbs D. The science of succes. The Atlantic Monthly, December, 60-68. Retrieved from: http://www.theatlantic.com/magazine/archive/2009/12/the-science-of-success/7761/. 2009.

56. Belsky J. Theory testing, effect-size evaluation, and differential susceptibility to rearing influence: the case of mothering and attachment. Child Dev. 1997 Aug;68(4):598-600.

57. Suomi SJ. Early determinants of behaviour: evidence from primate studies. Br Med Bull. 1997 Jan;53(1):170-84.

58. van Zeijl J, Mesman J, Stolk MN, et al. Differential susceptibility to discipline: the moderating effect of child temperament on the association between maternal discipline and early childhood externalizing problems. J Fam Psychol. 2007 Dec;21(4):626-36.

59. Kochanska G, Aksan N, Joy ME. Children's fearfulness as a moderator of parenting in early socialization: Two longitudinal studies. Dev Psychol. 2007 Jan;43(1):222-37.

60. Goedhart G, Vrijkotte TG, Roseboom TJ, van der Wal MF, Cuijpers P, Bonsel GJ. Maternal cortisol and offspring birthweight: results from a large prospective cohort study. Psychoneuroendocrinology. 2010 Jun;35(5):644-52.

61. Kivlighan KT, DiPietro JA, Costigan KA, Laudenslager ML. Diurnal rhythm of cortisol during late pregnancy: associations with maternal psychological well-being and fetal growth. Psychoneuroendocrinology. 2008 Oct;33(9): 1225-35.

62. Economides DL, Nicolaides KH, Linton EA, Perry LA, Chard T. Plasma cortisol and adrenocorticotropin in appropriate and small for gestational age fetuses. Fetal Ther. 1988;3(3):158-64.

63. Phillips DI, Walker BR, Reynolds RM, et al. Low birth weight predicts elevated plasma cortisol concentrations in adults from 3 populations. Hypertension. 2000 Jun;35(6):1301-6.

64. Levitt NS, Lambert EV, Woods D, Hales CN, Andrew R, Seckl JR. Impaired glucose tolerance and elevated blood pressure in low birth weight, nonobese, young south african adults: early programming of cortisol axis. J Clin Endocrinol Metab. 2000 Dec;85(12):4611-8.

65. Reynolds RM, Walker BR, Syddall HE, et al. Altered control of cortisol secretion in adult men with low birth weight and cardiovascular risk factors. J Clin Endocrinol Metab. 2001 Jan;86(1):245-50.

66. Jones A, Godfrey KM, Wood P, Osmond C, Goulden P, Phillips DI. Fetal growth and the adrenocortical response to psychological stress. J Clin Endocrinol Metab. 2006 May;91(5):1868-71.

67. Wust S, Entringer S, Federenko IS, Schlotz W, Hellhammer DH. Birth weight is associated with salivary cortisol responses to psychosocial stress in adult life. Psychoneuroendocrinology. 2005 Jul;30(6):591-8.

68. (NELP) NELP. Developing early literacy: Report of the National Early Literacy Panel. Washington, DC: National Institute for Literacy. Available at http://www.nifl.gov/earlychildhood/NELP/NELPreport.html. 2008.