ISBN 978-82-326-2060-9 (printed ver.)
ISBN 978-82-326-2061-6 (electronic ver.)
ISSN 1503-8181

Tormod Rogne

Doctoral thesis

Doctoral theses at NTNU, 2016:358

Tormod Rogne

Causes and Consequences of Small Size at Birth

Maternal Glucose Tolerance and Vitamin B12 Levels in Pregnancy, and Later Brain Volumes and Cognitive Function in the Offspring

Doctoral theses at NTNU, 2016:358

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of Philosophiae Doctor
Faculty of Medicine
Department of Public Health and General Practice

NTNU
Norwegian University of Science and Technology

NTNU

NTNU
Norwegian University of Science and Technology

Tormod Rogne

Causes and Consequences of Small Size at Birth

Maternal Glucose Tolerance and Vitamin B12 Levels in Pregnancy, and Later Brain Volumes and Cognitive Function in the Offspring

Thesis for the Degree of Philosophiae Doctor

Trondheim, December 2016

Norwegian University of Science and Technology
Faculty of Medicine
Department of Public Health and General Practice

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Medicine
Department of Public Health and General Practice

ISBN 978-82-326-2060-9 (printed ver.)
ISBN 978-82-326-2061-6 (electronic ver.)
ISSN 1503-8181

Doctoral theses at NTNU, 2016:358

Printed by NTNU Grafisk senter

Årsaker til og konsekvenser av liten fødselsstørrelse
Mors glukosetoleranse og vitamin B12-konsentrasjon i svangerskapet, og senere hjernevolumer og kognitive evner hos barnet

Liten fødselsstørrelse kan være forårsaket av dårlig vekst i fosterlivet eller tidlig fødsel. Lav fødselsvekt og for tidlig fødsel er på verdensbasis årsak til nær halvparten av alle barnedødsfall de første 28 dagene etter fødselen. I tillegg kan liten fødselsstørrelse ha konsekvenser senere i livet, for eksempel reduserte kognitive evner. Men når er et nyfødt barn *for* lite? Noen barn kan være født naturlig små (for eksempel fordi foreldrene også er små), mens andre barn er født mindre enn de skulle ha blitt. Sistnevnte betegnes ofte *hemmet fostervekst*. Målet med denne avhandlingen var å utforske noen årsaker til og konsekvenser av liten fødselsstørrelse. Vi studerte to potensielle årsaker til liten fødselsstørrelse: mors glukosetoleranse og nivået av vitamin B12 (B12) i mors blod i svangerskapet. Vi studerte også sammenhengen mellom vekstmønsteret i fosterlivet og barnets kognitive evner og hjernevolumer.

Første artikkel i avhandlingen var basert på data fra en skandinavisk observasjonsstudie av gravide kvinner. Veksten til fosteret ble beregnet ved hjelp av repeterte ultralydmålinger i andre og tredje trimester, og i tredje trimester ble også mors glukosetoleranse vurdert. Til sammen 855 kvinner ble inkludert i studien. Vi fant en tendens til redusert fostervekst hos gravide kvinner med høy glukosetoleranse. Disse kvinnene fødte også slankere barn, men med tilsvarende fødselsvekt, sammenliknet med barna født av de øvrige kvinnene.

Barn født av kvinnene i den skandinaviske observasjonsstudien ble fulgt opp etter fødselen. Barnas kognitive evner ble vurdert ved fem og ni års alder, og hjernevolumer ble målt ved 15 års alder. Andre artikkel inkluderte 83 barn født SGA til termin ("liten for svangerskapslengde" (engelsk "small-for-gestational-age")), det vil si de 10% med lavest fødselsvekt for sin svangerskapslengde, samt 105 ikke-SGA barn. Basert på ultralydmålinger i svangerskapet påviste vi hemmet fostervekst blant 13 barn i SGA-gruppen, mens 36 barn viste ingen tegn til hemmet fostervekst. Vi fant at barn født SGA på grunn av hemmet fostervekst – men ikke de som ble født SGA uten veksthemming – gjorde det dårligere på kognitive tester ved fem og ni år, og hadde mindre volum av hjernestrukturen thalamus og lillehjernens hvitsubstans, sammenliknet med barn født med normal fødselsvekt.

Siste artikkel i avhandlingen er en systematisk oversiktsartikkel hvor vi vurderte hvordan mors nivå av vitamin B12 i svangerskapet henger sammen med lengden på svangerskapet og barnets fødselsvekt. Kvalifiserte studier delte individuelle pasientdata. Tjueto kvalifiserte studier ble identifisert, hvorav 18 ble inkludert i meta-analysen (11,216 graviditeter, 94% av alle kvalifiserte graviditeter). Vi fant at B12-mangel hos gravide var knyttet til en økt risiko for tidlig fødsel og å føde barn med lav fødselsvekt, men ikke SGA-fødsel.

Denne avhandlingen finner at gravide kvinner med høy glukosetoleranse kan ha en økt risiko for å bære fostre med suboptimal vekst. B12-nivået i mors blod, derimot, ser ut til å være assosiert med lengden på svangerskapet, men ikke fostervekst. Risikoen for redusert kognitiv funksjon og mindre hjernevolumer i barndommen ser ut til å være større hos barn født små på grunn av hemmet fostervekst enn hos øvrige barn født små.

Navn kandidat: Tormod Rogne
Institutt: Institutt for samfunnsmedisin
Veiledere: Kari R. Risnes (hovedveileder), Geir W. Jacobsen (biveileder) og Marit P. Martinussen (biveileder)
Finansieringskilde: Norges teknisk-naturvitenskapelige universitet

Ovennevnte avhandling er funnet verdig til å forsvares offentlig for graden PhD i medisin. Disputas finner sted i auditoriet MTA, torsdag 15. desember 2016, kl 12:15.

Table of Contents

Acknowledgements

This thesis was carried out at the Department of Public Health and General Practice, Faculty of Medicine, Norwegian University of Science and Technology (NTNU). The Research Council of Norway funded the first part of this project through the Medical Student Research Program, while the Faculty of Medicine, NTNU, funded the last part through the PhD in Medicine Program. During my research period I was also enrolled in a visiting researcher program at Yale University, New Haven, CT, USA.

First, I would like to thank my main supervisor, Kari Risnes, whose energy and knowledge has been key to land this thesis. In our meetings at work or over meals abroad, Kari has always been cheerful, fair and friendly. I have also learned a lot from Kari's comprehensive understanding of pediatrics, systematic reviews and evidence-based medicine. In short, Kari's academic proficiency and affable character have allowed for important conversations and discussions on any topic.

Geir Jacobsen was my main supervisor when I was a medical student researcher, and has been my co-supervisor on the PhD thesis. I would like to thank Geir for his advice and assistance the last five years, for always being joyful and warm, for gesturing me into his office even before knocking, and for inspiring me with his enthusiasm for public health and perinatal epidemiology. Also, I would very much like to thank my co-supervisor Marit Martinussen for generously sharing her extensive knowledge of fetal growth and brain development, for helping write and review the second and third paper, and for her friendliness.

Jon Skranes and Heidi Furre Østgård co-authored the second paper, and I would like to thank them for promptly responding to all my enquires, for sharing their thoughts and opinions, and for reviewing and editing the second paper. I am very grateful to Eva Skovlund who advised how to run the statistics in the final paper, and helped edit that paper.

Thanks also to Myrte Tielemans in Rotterdam who greatly contributed in composing and editing the systematic review. In addition to providing insight on maternal nutrition in pregnancy and statistical analyses, I very much appreciate how Myrte welcomed me and Kari on our visits to Rotterdam, our lunch breaks and dinners, and our discussions on how best to furnish our new apartments. I am thankful to Oscar Franco who facilitated the cooperation between NTNU and Erasmus MC, Rotterdam, and for his motivation.

I would like to thank Andreas Aass-Engstrøm for working with me on the second paper, for the fruitful discussions concerning fetal growth and cognitive outcome, and for his light-heartedness. Trond Nordseth has been an important external advisor and I greatly value his encouragement. Thanks also to Marius Flasnes, Hilde Lauritzen and Ivar Følling for their critical input and assistance in the first paper. Anne-Mette Podhorny proofread this thesis on her vacation in Turkey, which all readers now greatly benefit from; thank you!

A special thanks to Michael and Maryann Bracken for welcoming me to New Haven, for teaching me how to shuck corn and brew American coffee, and for making my stay across the Atlantic so enjoyable. Michael's advise and expertise was essential to compose the systematic review in this thesis. Through his lectures at Yale and amusing anecdotes at Woodland Coffee & Tea, Michael has taught me most all I know about evidence-based medicine.

Last but not least, I would like to thank my friends and my family, for listening to all my praise and complaints about life as a researcher, and for taking my mind off work.

List of papers

Paper I Rogne T, Jacobsen GW. Association between low blood glucose increase during glucose tolerance tests in pregnancy and impaired fetal growth. Acta Obstet Gynecol Scand. 2014;93:1160-9.

Paper II Rogne T, Engstrøm AA, Jacobsen GW, Skranes J, Østgård HF, Martinussen M. Fetal growth, cognitive function, and brain volumes in childhood and adolescence. Obstet Gynecol. 2015;125:673-82.

Paper III Rogne T, Tielemans MJ, Chong MFF, Yajnik CS, Krishnaveni GV, Poston L, Jaddoe VWV, Steegers EAP, Joshi S, Chong YS, Godfrey KM, Yap FKP, Yahyaoui R, Thomas T, Hay G, Hogeveen M, Demir A, Saravanan P, Skovlund E, Martinussen MP, Jacobsen GW, Franco OH, Bracken MB, Risnes KR. Maternal vitamin B12 in pregnancy and risk of preterm birth and low birth weight: A systematic review and individual participant data meta-analysis. Am J Epidemiol (accepted for publication).

Acronyms

Δ:	Delta
AGA:	Appropriate-for-gestational-age
B12:	Vitamin B12
BMI:	Body mass index
BPD:	Biparietal diameter
CI:	Confidence interval
CoA:	Coenzyme A
CP:	Cerebral palsy
DM:	Overt diabetes mellitus
EBM:	Evidence-based medicine
FGR:	Fetal growth restriction
GDM:	Gestational diabetes mellitus
IGF:	Insulin-like growth factor
IPD:	Individual patient data
IQ:	Intelligence quotient
IUGR:	Intrauterine growth restriction
KRR:	Kari R Risnes
LBW:	Low birth weight
MJT:	Myrte J Tielemans
MRI:	Magnetic resonance imaging
NTNU:	Norwegian University of Science and Technology
OGTT:	Oral glucose tolerance test
PROM:	Premature rupture of membranes
RCT:	Randomized controlled trial
RR:	Risk ratio
SD:	Standard deviation
SDG:	Sustainable development goal
SGA:	Small-for-gestational-age
TR:	Tormod Rogne
WHO:	World Health Organization
WISC-R:	Wechsler Intelligence Scale for Children Revised
WPPSI-R:	Wechsler Preschool and Primary Scale of Intelligence Revised

Summary

Small size at birth may be caused by restricted fetal growth or preterm birth. Each year, preterm birth and low birth weight (LBW) cause roughly half of the 2.9 million neonatal deaths globally. Additionally, small birth size has been associated with long-term outcomes such as all-cause mortality and reduced cognitive function. But when is a newborn too small? Some newborns may be physiologically small (e.g. small parents), while others are born small due to a pathological process. The latter is often referred to as *fetal growth restriction* (FGR). The aim of this thesis was to explore some potential causes and consequences of small birth size. We studied two potential causes of small birth size: maternal glucose tolerance and vitamin B12 (B12) levels in pregnancy. We also studied the association between fetal growth pattern and offspring cognitive function and regional brain volumes.

The first paper in this thesis was based on a large, long-term follow-up study in Scandinavia. The population was enriched with women at an increased risk of giving birth to a child with LBW. Ultrasound measurements were used to estimate fetal growth in the second and third trimesters, and an oral glucose tolerance test was performed in the third trimester. The difference between the two-hour and fasting blood glucose values was labeled delta (Δ) glucose. A total of 855 women were included in the study. We found that the most glucose tolerant women, identified by a low Δ glucose, were associated with an increased risk of carrying fetuses with suboptimal growth. These women also gave birth to thinner newborns, but of similar weight, compared with the other women.

Offspring of the women followed in the Scandinavian study were followed after birth. Cognitive function was assessed at five and nine years of age, and regional brain volumes were estimated at age 15 years. The second paper included 83 children born small-for-gestational-age (SGA; birth weight $<10^{th}$ percentile) at term and 105 non-SGA children. Based on serial ultrasound measurements, 13 children in the SGA-group were classified as FGR (SGA-FGR)

and 36 were classified as non-FGR (SGA non-FGR). We found that children born SGA due to FGR (SGA-FGR) – but not those born constitutionally small (SGA non-FGR) – had impaired performance intelligence quotient scores and smaller thalamic and cerebellar white matter volumes compared with controls.

The last paper in this thesis is a systematic review where we evaluated the association between maternal vitamin B12 blood levels in pregnancy and newborn birth weight and length of gestation. Eligible studies provided individual patient data (IPD), and when IPD could not be provided, relevant estimates from individual studies were included in the analyses. Twenty-two eligible studies were identified, of which 18 were included in the meta-analysis (11,216 pregnancies; 94% of all eligible pregnancies). B12-deficiency in pregnancy was associated with an increased risk of preterm birth and LBW, but not SGA. The increased risk of preterm birth among B12-deficient women was similar in high-income countries and low- and middle-income countries.

This thesis suggests that pregnant women who have a high glucose tolerance may have an increased risk for carrying fetuses with suboptimal growth. Vitamin B12, however, seems to be associated with the length of gestation, but not fetal growth. The risk of reduced cognitive function and smaller regional brain volumes in childhood and adolescence seems to be higher among children born small due to restricted fetal growth than in other children born small.

1. Introduction

1.1 Perspective

"We are beginning to identify processes that link fetal and infant growth with cardiovascular disease. […] Its [fetal growth failure] causes are unknown, but maternal nutrition is an obvious suspect."[1]

David Barker, 1990

With what has later been coined the "fetal origins hypothesis", Barker linked the fetal milieu and birth weight with later chronic diseases. Weight at birth is a result of how well the fetus has grown, and for how long. Birth weight is the common denominator in this thesis, and is both a sensitive marker of maternal and fetal health during pregnancy,[2] and an important predictor of later morbidity and mortality.[3] This thesis will cover examples of both aspects (Figure 1.1).

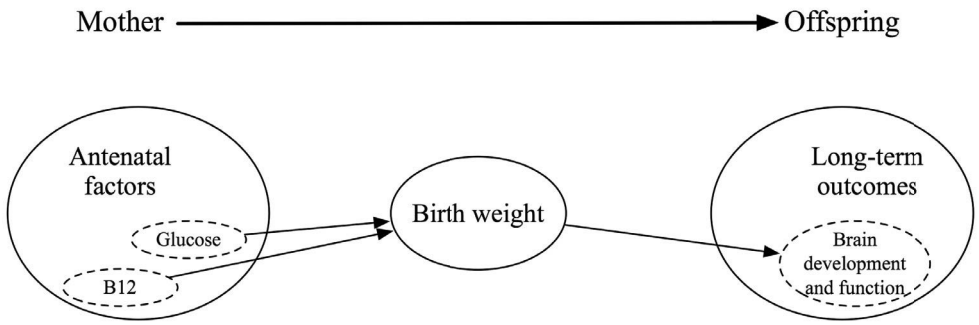


Figure 1.1 Perspective of thesis. The association between glucose and birth weight, and vitamin B12 and birth weight, will be addressed in *Paper I* and *Paper III*, respectively. Birth weight in relation to brain volumes and cognitive function will be addressed in *Paper II*.

Maternal nutrition prior to and during pregnancy is important for maternal, fetal, and child health outcomes, including birth weight.[4] Nutrient intake prior to and during pregnancy may directly affect both the fetal growth pattern and the length of gestation.[4] In terms of fetal well-

being, maternal nutrient *intake* is not equivalent to *fetal availability* of nutrients. For instance, reduced placenta function may lead to poor availability of nutrients to the fetus despite high intake.[5] While "nutrition" encompasses a vast range of constituents, e.g. fatty acids, carbohydrates and micronutrients, this thesis will focus on two nutrients; glucose and vitamin B12 (B12; Figure 1.1). The former has been extensively addressed in the setting of increased availability to the fetus related to metabolic disorders in the mother, such as gestational diabetes mellitus (GDM).[6] Low glucose levels in the mother could potentially influence fetal nutrition, but this has been less evaluated. The association between low blood glucose and fetal growth is the topic in *Paper I.* Scarcity of vitamin B12 is common during pregnancy:[7] Its consequences in terms of fetal growth and length of gestation, though, are unknown, and were addressed in *Paper III*.

Birth weight is a key determinant of later morbidity and mortality.[8] While low birth weight (LBW) is strongly correlated with later adverse outcomes on a population level, the correlation is not that strong on an individual level.[8] Two newborns of the same weight may have very different risks of later chronic diseases. This could potentially be explained by factors contributing to the weight at birth. A more refined definition of suboptimal birth weight may be warranted. This was explored in *Paper II*, with cognitive function and brain volumes in childhood and adolescence as the outcomes of interest (Figure 1.1).

The research from which we generate new hypothesis and make clinical decisions vary in both design and reliability. The ideal setting to draw conclusions on causality is a trial where patients are randomized to receive either a new treatment or the current best treatment. However, this design may be inappropriate in many settings, e.g. smoking exposure in relation to lung cancer. In these settings, observational studies are warranted. Furthermore, as it is not uncommon – and indeed implicit in the framework of falsifiability in scientific research – that individual studies are contradicting, synthesis of the state of current knowledge is important.

Well conducted syntheses are arguably the best source of evidence.[9] The understanding of how to conduct and interpret different types of epidemiological studies is becoming increasingly important for the public health researcher. This thesis is based on two observational studies, *Papers I* and *II*, and a systematic overview of multiple observational studies, *Paper III*.

1.2 Context

There are 15 million preterm births and 20 million LBW births globally each year:[10] In Norway, in 2014, the respective numbers were 3,294 (5.6%) and 2,914 (4.9%).[11] The greatest burden of newborns born small is found in South Asia (up to 40%), while rates of preterm births are highest in Africa (up to 18%).[10] Of the 2.9 million neonatal deaths globally, 1 million deaths are estimated to be directly caused by preterm birth, and an additional 0.5 million indirectly caused by preterm birth (mainly infections).[10]

The Millennium Development Goals were just concluded, and were aimed at reducing the under-five year mortality rate by two thirds from 1990 to 2015.[12] We are now entering the era of the Sustainable Development Goals (SDGs), with the ambition of further reducing the under-five year mortality from 43 per 1,000 live births in 2015 to 25 per 1,000 live births in 2030. The UN Inter-agency Group for Child Mortality Estimation estimates that the proportion of under-five year deaths constituted by neonatal deaths will increase from 45% in 2015 to 52% in 2030.[12]

As preterm birth and LBW cause half of all neonatal deaths worldwide, prevention of these pregnancy outcomes is key in order to reduce neonatal mortality and reach the SDGs. Low newborn weight is also important for later all-cause mortality.[3] Still, it is unclear whether it is the birth weight *per se* or the fetal growth pattern leading up to the final weight which is important.

1.3 Evidence-based medicine

The eager medical student is early on faced with details about frequencies of diseases and the effects of specific drugs; numbers and statements perceived as facts. After the introduction of confidence intervals somewhat later in the education, the "truth" is challenged with a sense of uncertainty. Alas, with increasing knowledge and experience on the rational, conduct and interpretation of medical research, the once comprehensible science of medicine has been thrown into disarray. When the student stumbles upon a quote by the influential epidemiologist J. Ioannidis that "it can be proven that most claimed research findings are false",[13] our physician-to-be is already in deep thoughts about an alternative career as mathematician.

Evidence-based medicine (EBM) is "the integration of the best research evidence with our clinical expertise and our patient's unique values and circumstances".[14] EBM also includes guidelines and health policy decision-making. The practice of EBM has been divided into several steps: asking the right question, finding the evidence, appraising the evidence, integrating the evidence into clinical practice, and evaluating its effectiveness.[15] However intuitive it may sound that decisions should be "evidence-based", there are discrepant opinions of what "evidence" is. There are many sources of information, and learning how to critically read studies is becoming increasingly important for clinicians.[16] Even Norwegian physicians have a way to go to become proficient in EBM.[17]

As will be discussed in this section, most research may be wasted, and there is often a gap between the evidence at hand and how medicine is practiced. Encouragingly, efforts are being made to solve these issues. One of the proposed solutions is the conduct of systematic reviews such as *Paper III* of this thesis, and will be discussed in more detail.

1.3.1 Research waste

Discouragingly, it has been estimated that 85% of medical research may be wasted.[18] Some of the factors contributing to this waste of research are that the wrong research questions are being addressed (e.g. not patient-centered), that the design of studies is poor, that the reporting of findings is inadequate or that findings are not published at all, and that what is published may be biased (e.g. conflicting interests).[18]

There are many proposed solutions to avoid research waste and to improve evidence-base in medicine. A series in *The Lancet* in 2014 put forward many recommendations: One should ensure that information of research in progress is readily available;[9] the standards of study design should be increased;[19] one should ensure integration of research in everyday clinical practice;[20] facilitate sharing of individual patient data (IPD);[21] and improve the reporting of studies.[22]

One of the recommendations highlighted in the series was that research funders and regulators should demand that when research priorities are set, they should be justified by systematic reviews.[9] A systematic review is often described as the most reliable study design (Figure 1.2). Many funders will not support trials unless there is evidence from systematic reviews that a trial is called for.[23] Ideally, conduct of systematic reviews should be incorporated into the training of research fellows.[18,24] Additionally, systematic reviews play an integral role in closing the gap between research evidence and clinical practice.[25] Systematic reviews will be further discussed in the following section.

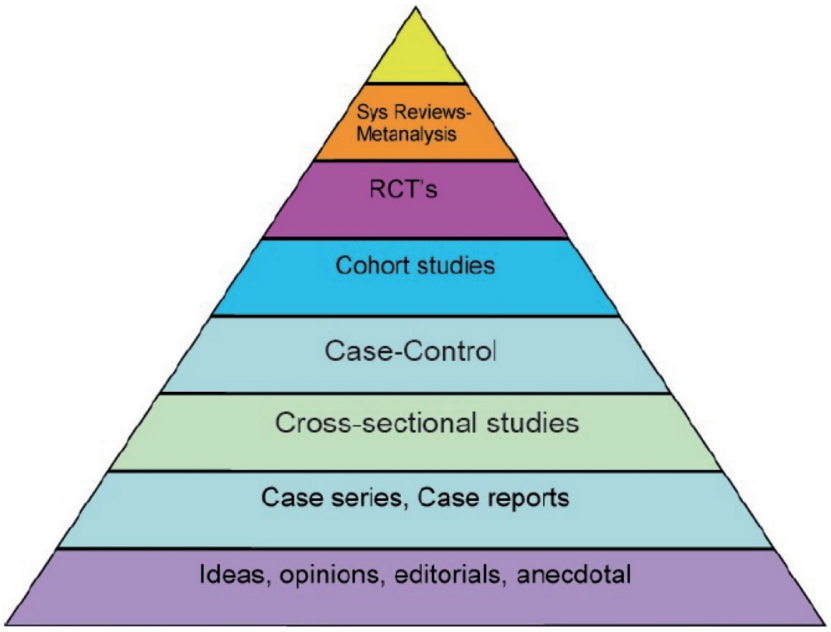


Figure 1.2 Evidence-based medicine pyramid, illustrating the quality of evidence by study design. RCT, randomized controlled trial. Illustration: blogs.bmj.com.

1.3.2 Systematic reviews

At its core, systematic reviews seek to identify all available evidence on a specific clinical problem, trying to make an informed decision as to what is the "typical" finding. It may be defined as "a review of a clearly formulated question that uses systematic and explicit methods to identify, select and critically appraise relevant research, and to collect and analyze data from studies that are included in the review."[26] A detailed protocol should be followed and made publicly available.

In order to analyze the data one may conduct a *meta-analysis*. Sometimes misinterpreted as a synonym for a systematic review, a meta-analysis is merely a statistical technique used to pool results from individual studies. As discussed below, it may often be wise to refrain from including meta-analyses in systematic reviews.

The notion to summarize published research dates back many centuries; one example being the *Account of the foxglove* by William Withering in 1785, on the use of digitalis for

treating heart disease. Combination of data from several studies was first introduced to the medical sciences by Karl Pearson in 1904.[27] It took many decades, though, before the why's and how's of systematic reviews were seriously addressed. A great step forwards in EBM was made in the mid-1970s when a group of researchers based in Oxford sought to evaluate interventions related to pregnancy and childbirth. Published and unpublished data on 600 interventions were collected, and in 1989 the group published their findings in the book *Effective Care in Pregnancy and Childbirth*.[28] This was the start of *The Cochrane Collaboration,* one of the most influential sources of EBM.

One of the founders of the Cochrane Collaboration, Sir Iain Chalmers, argues that all research should be preceded by systematic reviews.[9] But why are systematic reviews so significant? Some of the most important reasons are addressed below, as are some caveats.

Research question already covered

Although both funders of research and journal editors give an emphasis to the novelty of research, there are many examples of how existing evidence has been ignored. A systematic review and meta-analysis of the use of streptokinase for acute infarction illustrates this perfectly:[29] The review authors found that after eight trials, involving 2,432 patients, enough evidence of the effect of streptokinase had been gathered. However, in the 15 years following the eighth trial, an additional 25 trials were conducted, involving another 34,542 patients, resulting only in a somewhat more precise estimate. Thousands of volunteers were thus unnecessarily randomized to receive placebo treatment, depriving them of the opportunity to receive medication documented to reduce mortality. Systematic reviews should be able to identify all existing evidence on a topic, concluding whether more research should be carried out or not.

All in one place

Given the rate by which new research is published, it has become increasingly difficult for researchers and clinicians to keep up with new evidence.[23] Additionally, the research receiving most media attention,[30] or being most cited in academic journals,[31] does not necessarily represent the state of current knowledge. Narrative reviews, i.e. reviews where the authors try to summarize the literature without the strict protocol and criteria followed in systematic reviews, have been the traditional way to synopsize the evidence within a subject. Such reviews may be of value if proceeded with care, but they are often subject to both publication and reporting bias,[32] along with a "file drawer" problem (i.e. ignorance of null-findings).[33] By design, systematic reviews should evaluate all studies on the research question of interest. Ideally, also unpublished studies should be sought out. Arguably, systematic reviews provide a more objective overview of the literature.

With great statistical power comes great epidemiological responsibility

With more subjects under study, systematic reviews have increased power to identify effects and associations otherwise discarded due to high levels of uncertainty. In epidemiological terms, the probability of false negative results is reduced. In randomized controlled trials the only baseline difference between the intervention group and the comparison group is due to random error (ideally). Varying incidence of the outcome between the groups may therefore reasonably be ascribed to the intervention. Systematic review of such studies will provide precise, unbiased estimates (Figure 1.3). Pooling of results from observational studies will also provide precise estimates, but the results may be misleading. For instance, separately meta-analyzing results from observational studies and randomized controlled trials of the association between beta-carotene intake and cardiovascular mortality, the former analysis found a risk reduction from intake of beta-carotene, while the latter found an increase in risk.[34] This

difference was probably due to confounding and bias, which one should always suspect in non-randomized studies. Recently, 235 types of biases were identified in biomedical research.[35] One of the biases identified was recall bias, and was probably at play in the following example: A meta-analysis based on retrospective observational studies found a 30% increased risk of breast cancer among women with a history of induced abortions.[36] In contrast, a register study of 1.5 million women found no association between induced abortions and later breast cancer.[37] The latter study did not permit selective memory and feeling of guilt to influence the statistics. For this reason, critically appraising the individual studies is one of the key tasks when conducting systematic reviews. When there is suspicion of bias and unaccounted confounding, one should refrain from including meta-analyses in systematic reviews.[34]

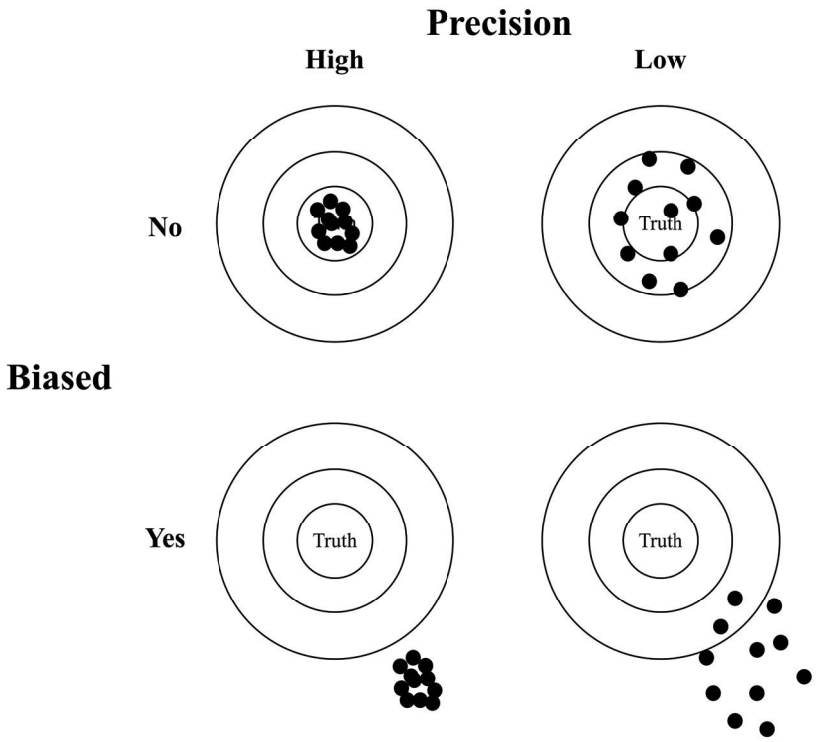


Figure 1.3 "Epidemiological dartboard" of the difference between precision and bias. Each dot represents a study; the center of the board represents the true effect under study. Meta-analyses of large RCTs will tend to be in the top left corner. The bottom left corner presents a scenario where a meta-analysis has provided very precise, but biased results, typically based on poorly designed observational studies. Adapted from the course *CDE 650 Introduction to Evidence-Based Medicine and Health Care* with permission from M B Bracken, Yale Center for Perinatal, Pediatric and Environmental Epidemiology, Yale University.

Hypothesis generating

With a larger sample size, there is also an opportunity to explore subgroup effects and generate new hypotheses. Topics that have not been sufficiently researched will also be identified from systematic reviews.

Publication bias and selective reporting bias

A considerable challenge in the medical sciences is the selective reporting of research findings. One problem is the incomplete or lack of reporting of the outcomes under study, known as *selective reporting bias*.[38] Other studies are not published at all, i.e. *publication bias*.[38] Additionally, compared with studies with negative findings, studies with statistically significant findings are more likely to be published in English and are more likely to be cited, which make these studies easier to be identified.[39] These forms of bias may distort the literature, and they are especially important in the setting of systematic reviews. In terms of trials on the efficacy of drugs, there is evidence to support that studies showing a positive effect has a greater tendency to be published than studies with a null-finding or negative effect.[40] The same is probably true for observational studies.[41] Conflicting interest with regards to "desired" results may further add bias. To curb publication bias of trials, it is now often required that trials must be registered.[42] There has been a call for similar registries for observational studies.[41,43] The reporting of outcomes in published reports are also subject to selection, often by leaving out null-findings.[40] To ensure proper reporting, several guidelines have been constructed (e.g. the PRISMA statement for systematic reviews[44] and the STROBE guideline for observational studies).[45] Many journals require both preregistration of trials and that reporting guidelines have been followed (e.g. The Lancet). In order to minimize bias introduced by selective reporting and publication, review authors are encouraged to retrieve data from unpublished studies, along with results not reported in the published reports. The gold standard of systematic reviews is

one including an IPD meta-analysis.[46] In an IPD meta-analysis, the authors will collect raw data from the studies included in the review. This permits an increased opportunity to conduct the desired analyses, adjust for confounding, and explore subgroup effects and heterogeneity.[47] Hence, more studies may be included in the meta-analysis and the statistical between-study heterogeneity is greatly reduced. The advantages of IPD meta-analyses compared with traditional meta-analyses in the setting of B12-deficiency in pregnancy is readily illustrated in section *5.4.2 Systematic review*.

As have been discussed, (properly conducted) systematic reviews are important to reduce waste in medical research. Additionally, clinical practice should be based on high quality systematic reviews. The ideal chain of events from exploring a hypothesis to implementing it into clinical practice would be the following: High quality randomized controlled trial (RCT) → rapid replication in a few additional RCTs → rapid, high quality systematic review (ideally with IPD meta-analysis) → rapid, high quality clinical guidelines → rapid widespread adoption in practice (M B Bracken, Yale University, personal communication, 2013). Unfortunately, there is evidence to support that clinical guidelines often fail to identify good quality evidence.[48] Additionally, not all topics may be evaluated through the suggested chain of events (e.g. the use of parachutes to prevent death related to gravitational challenge).[49] Discouragingly, research show that 2/3 of guidelines are not followed in clinical practice.[50]

1.4 Definitions on newborn size

Birth weight has been a popular outcome and exposure measure in epidemiological research for decades. One of the reasons for this is that birth weight is routinely recorded in clinical practice, leaving large numbers of relatively accurate records. Additionally, birth weight has been demonstrated to be an important predictor of both infant and adult morbidity and

mortality.[8] The final weight of the newborn is a result of how well the fetus has grown (i.e. the fetal growth pattern), and for how long (i.e. gestational age at birth).

Birth weight has previously been regarded as the single most important factor for perinatal mortality. One study reported it to explain 90% of the variance, while length of gestation was found to explain a mere 5%.[51] Even so, given that birth weight is so heavily dependent on gestational age, it may be difficult to disentangle the two. When separating *relative birth weights* for any given gestational age from gestational age at delivery, Wilcox et al. found that gestational age independently could explain much of the perinatal mortality related to birth weight as a crude measure.[52] Now, reduction of preterm birth (gestational age at delivery less than 37 weeks) is regarded as one of the most important tasks in order to reduce neonatal and under-five year mortality.[10]

Traditionally, birth weight has often been dichotomized into LBW, defined as newborn weight below 2,500 g, and normal birth weight otherwise.[53] The rationale underlying this division is the assumption that some newborns are ill at birth, and others are healthy. The association between LBW and infant mortality is so strong that one may be tempted to assume a causal relationship. However, this may not be true. For instance, due to the high altitude in Colorado, the incidence of LBW is greater than what is observed in the rest of the USA; The infant mortality, though, is not increased.[8]

As have already been discussed, birth weight is a composite measure of both fetal growth and length of gestation. LBW may for this reason be a result of preterm birth, and both LBW and preterm birth are independent predictors of perinatal mortality.[52,54] Another definition has emerged to better serve as an indicator of restricted fetal growth: *Small-for-gestational-age* (SGA).

SGA is most commonly defined as the lowest 10 percentile of birth weight for a given gestational age, and is often classified by parity and sex as well.[55] As the SGA definition is

specific for the duration of pregnancy, it has been considered to be more accurate than LBW for identifying newborns at risk for later adverse health events.[56] Still, this definition is unable to clearly separate the pathologically small newborns from those physiologically small. For instance, among South Asian births, 10% will by definition be an SGA birth using South Asian reference charts, but when using Canadian ones, 21% will be classified as SGA.[57] One may apply stricter criteria for who to categorize as SGA, e.g. using the 3rd percentile rather than the 10th. Stricter criteria will be more specific for pathology, but may be less sensitive.

As early as in 1983, Wilcox argued that it was necessary to distinguish between newborns born SGA due to biological variability and those born SGA due to restricted growth.[58] The true, unmeasurable process causing birth of a newborn that is too small (in relative terms) is often labelled *fetal growth restriction* (FGR), alternatively *intrauterine growth restriction* (IUGR). Customized birth weight standards (e.g. accounting for maternal weight and height) have been argued to outperform population-based birth weight standards,[59] but there is lack of valid evidence for this assumption.[60] As placenta dysfunction is one of the major causes of FGR,[61] attempts have been made to monitor placenta function in order to identify FGR. Concern has been raised as to whether placenta function is a good predictor of FGR or not.[62,63] Perhaps the most accepted way to identify FGR is by directly following the fetal growth during pregnancy by use of serial ultrasound measurements.[64–66]

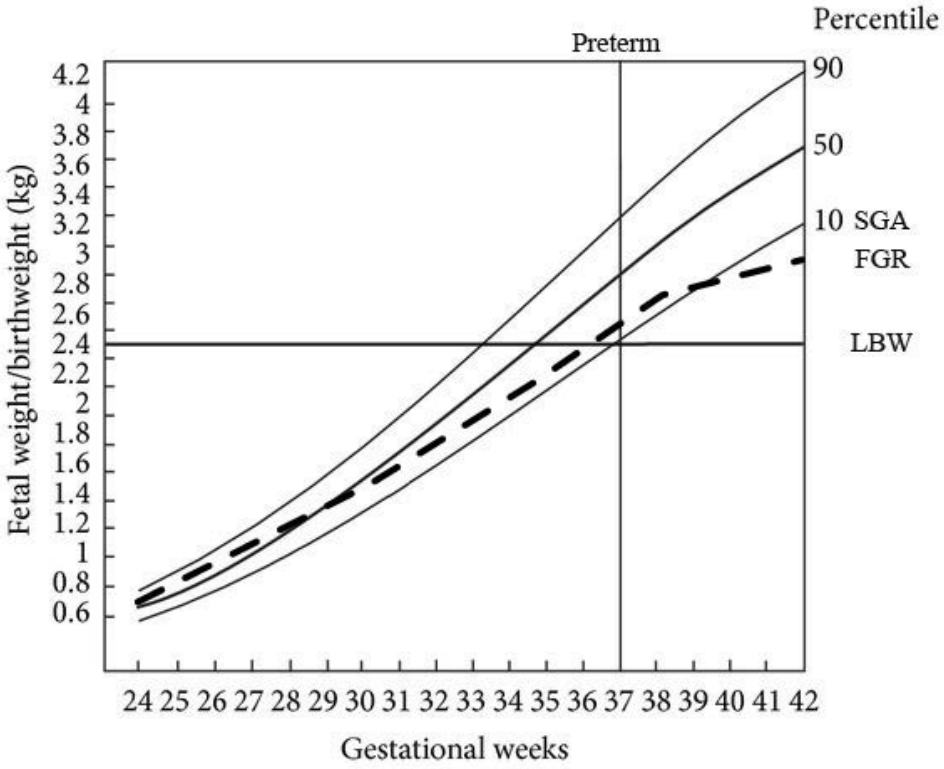


Figure 1.4 Fetal growth chart. Dashed line illustrates an example of restricted fetal growth. FGR, fetal growth restriction; LBW, low birth weight; SGA, small-for-gestational-age. Adapted from Haram et al.[67]

Figure 1.4 illustrates the four terms related to newborn size; LBW, SGA, FGR and preterm birth. Evident from the fetal growth chart, newborn size increases with increasing length of gestation. Preterm birth (solid vertical line) is therefore often related to LBW (solid horizontal line), but not necessarily SGA (the 10th percentile). An example of a restricted fetal growth pattern is illustrated by the dashed line: In early gestation the fetus has an estimated fetal weight above the median (the 50th percentile), followed by reduced growth in the third trimester. As demonstrated, the final weight of the FGR newborn may be well above the cut-off for LBW.

1.5 Causes of small birth size

The etiology of small size at birth is multifactorial and may be categorized into three groups: maternal, placental and fetal causes.[55,68–70] The role of individual risk factors in determining birth size is often not clear, and many etiological factors remain unknown.

It has been estimated that 15% to 20% of FGR cases are due to fetal factors.[68] They include infections, congenital malformations, and multiple gestation. The same risk factors are associated with preterm birth.[70] FGR may also cause preterm birth, and is a common reason for induced preterm delivery.[71,72]

Adequate placenta function is vital in order to supply the fetus with ample amounts of oxygen and nutrients. FGR is often caused by an abnormal development of the placenta, insufficient implantation, or placental villi dysfunction.[68,69]

The maternal risk factors for small birth size may be divided into demographic and environmental factors, and maternal diseases.[68–70] The former group constitutes risk factors such as previous LBW infant, short stature, young or old age (below 17 or above 34 years), and low socioeconomic status. One of the most important causes of FGR is maternal chronic vascular disease, mostly secondary to e.g. renal disease or hypertension, and in particular when these conditions appear in combination with preeclampsia.[69] Preeclampsia is also one of the most important reasons for induced preterm delivery.[71] In up to half of cases, the precursor of preterm birth is unidentified, but demographic factors and family history are important.[73] Environmental factors, e.g. cigarette smoking and substance abuse, further add to a detrimental fetal environment and may cause both FGR and preterm delivery.

The intrauterine environment, rather than genetics, seems to be the main determinant of size at birth. This has been demonstrated in a study that evaluated factors affecting newborn birth weight in pregnancies following ovum donation.[74] They found that while recipient's weight influenced the offspring birth weight, donor weight and birth weight of the donor's own

children did not correlate with birth weight of the offspring. Another study found fetal genes to explain 31% of the variation in birth weight, but only 11% of the variation in length of gestation.[75] Fetal substrate availability is probably key in explaining the importance of the intrauterine milieu. There is a large number of experimental studies on animals demonstrating the deleterious effects on fetal growth after reducing the availability of selected nutrients.[5] In "natural experiments" on humans, such as the Dutch Hunger Winter, one has observed reduced birth weight following restricted nutrient intake.[76]

There has been extensive research and public interest on the association between micronutrients and birth weight. Barker communicated that the natural suspect for small birth size was nutrition, and that optimal maternal nutrition was key in order to reduce adult chronic diseases of the offspring.[77] This unique opportunity for interventions to prevent adverse pregnancy outcomes makes micronutrients an interesting topic. A recent Cochrane review of trials found that multiple micronutrient supplementation in low- or middle-income countries reduced the risk of LBW and SGA births.[78] There was, however, no reduction in risk of preterm birth or neonatal mortality. Two micronutrients of increasing public health interest are glucose and vitamin B12. They have been studied as parts of the present thesis and will be discussed in more detail.

1.5.1 Glucose

A plethora of hormonal and metabolic changes take place in pregnancy. One that is important for fetal growth is alterations in the metabolism and distribution of glucose. With increasing gestational age, there is an increase in insulin resistance.[79] A consequence of this diabetogenic adaptation is an increased glucose concentration in the maternal bloodstream. Glucose is the most important nutrient for the growing fetus and the main determinant for fetal glucose concentration is that of its mother.[80,81] Since insulin does not cross the placenta, the fetus is

solely dependent on its own pancreatic production of insulin, of which glucose is the strongest stimulant (Figure 1.5).[82] Insulin is a key fetal growth factor, both directly and through stimulation of the insulin-like growth factor (IGF) axis.[83,84] For instance, it has been hypothesized that one reason why female newborns weigh less than male newborns is that female fetuses are more resistant to the trophic effects of insulin.[85] In short, the increased maternal insulin resistance in pregnancy aids the growing fetus in two very important ways: (1) Increased supply of glucose (energy) and (2) stimulation of fetal insulin production (growth).

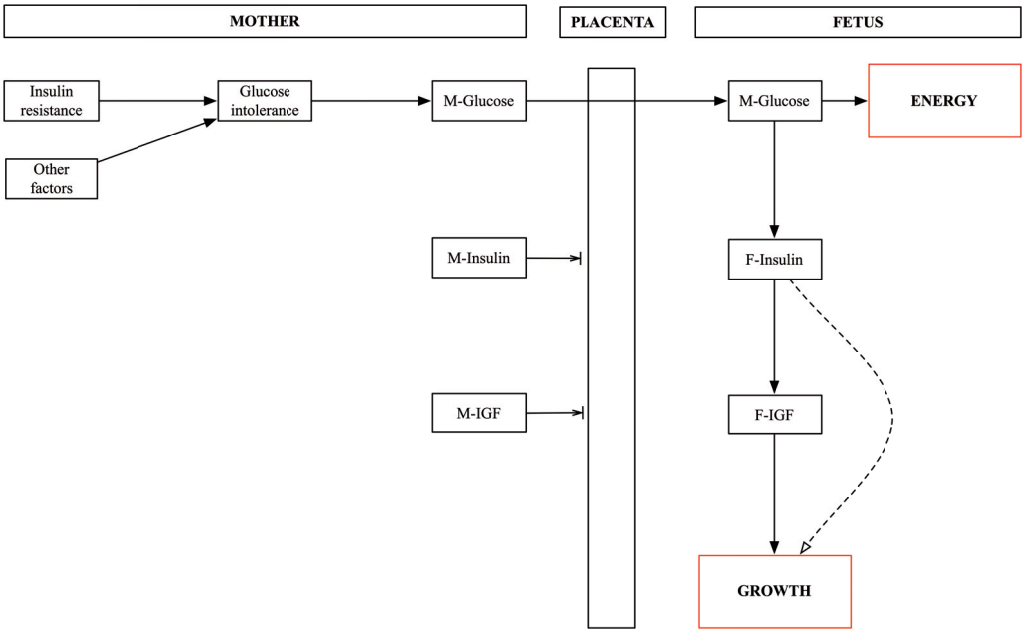


Figure 1.5 Schematic presentation of maternal glucose in relation to fetal growth, based on the text from Scholl et al.,[80] Oliver et al.,[83] Persson et al.,[81] and Lockitch.[82] Prefixes M and F for maternal and fetal, respectively. IGF, insulin-like growth factor.

Excessive insulin resistance, has been thoroughly investigated. It may lead to GDM or diabetes mellitus (DM).[6] Pregnancies affected by these conditions have increased levels of maternal blood glucose and insulin.[82] Fetal overgrowth and increased risk of macrosomia are some of the consequences.[86] In addition to fetal overgrowth, GDM or DM is also associated with

restricted fetal growth, especially among women with islet beta-cell autoantibodies.[87] Insulin resistance is also associated with an increased risk of preterm birth, perhaps partly mediated through preeclampsia.[88–90]

Far fewer studies have been conducted on women at the other end of the insulin resistance spectrum, namely those that are too insulin sensitive. There is, however, reason to suspect an impairment of fetal growth when the pregnant woman experiences suboptimal increase in insulin resistance.[91–93]

In today's antenatal care the primary reason for performing a routine oral glucose tolerance test (OGTT) is to diagnose and treat GDM and DM.[94] One might argue that too little attention has been paid to the women with inadequate rise in insulin resistance.

There are several ways to assess glucose and insulin metabolism. The "gold standard" for assessing insulin sensitivity is the hyperinsulinemic euglycemic glucose clamp technique.[95] Because this test is complex, e.g. involving several hours of continuous insulin infusion, simpler tests that indirectly measure insulin sensitivity have been developed. Most of these techniques involve measurement of both glucose and insulin at a fasting state, and preferably also after a glucose challenge.[95] The OGTT is a simple and widely applied test that evaluates glucose tolerance. After an overnight fast, a 75 g oral glucose load is given and fasting and two-hour capillary blood glucose values are recorded.[6] The test evaluates the body's ability to dispose of glucose after a meal. Glucose tolerance and insulin sensitivity are similar, but not equivalent. How the body metabolizes glucose depends on other factors in addition to the metabolic effects of insulin, such as secretion of insulin and the incretin effect.[95] The OGTT is thus a good test for glucose tolerance, but less so for insulin sensitivity.

For the purpose of *Paper I*, we assumed that the difference between two-hour blood glucose and fasting blood glucose from an OGTT, labelled delta (Δ) glucose, may potentially serve as a proxy for glucose tolerance and, to a lesser extent, insulin resistance. In a situation

with increased glucose *intolerance*, one would expect the Δ glucose to be *high*, since there is decreased uptake of glucose. And the opposite would potentially be the situation for increased glucose *tolerance.*

1.5.2 Vitamin B12

Vitamin B12, or cobalamin, plays an important role in a network of biochemical pathways called the one-carbon metabolism that donate and regenerate one-carbon units, such as the methyl group (Figure 1.6).[96] B12 is required as a cofactor to generate methionine from homocysteine.[97] Consequently, a lack of B12 may cause accumulation of homocysteine. By addition of adenosine triphosphate, methionine is further converted into S-adenosyl methionine, which in turn is an important methyl group donor in the methylation of DNA.[98] Importantly, B12 therefore contributes in the epigenetic mechanisms for control of gene expression.[96] The synthesis of DNA is also influenced by B12.[97] Another pathway in which B12 takes part is the formation of succinyl-coenzyme A (CoA), which is important for the cell's energy production.[99] Based on these biologically important properties, B12 is important for gene expression and cell growth, and it has been hypothesized that B12 may affect fetal growth.[100]

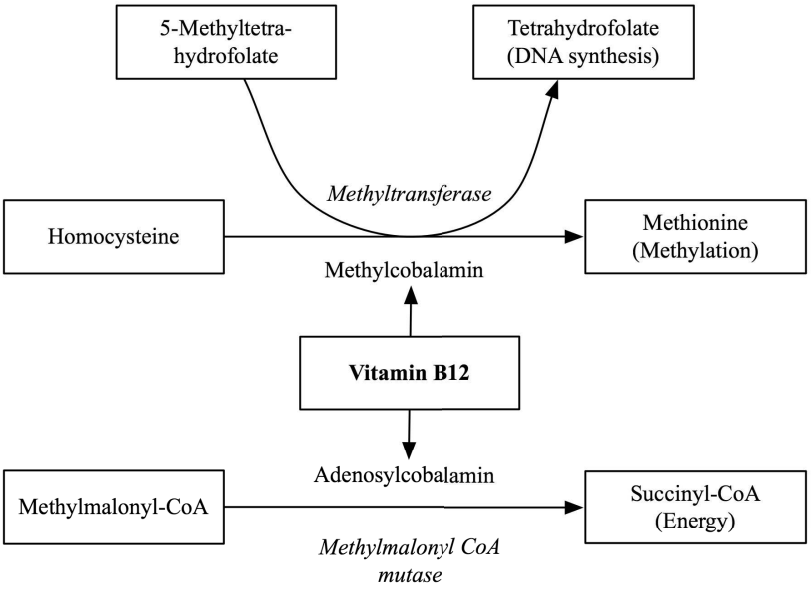


Figure 1.6 Biochemical pathways of vitamin B12, based on the text from Rush et al.,[96] Allen et al.,[97] Chiang et al.,[98] and Halarnkar et al.[99] Enzymes in italics. CoA, coenzyme A.

There is great range in the prevalence of B12-deficiency in pregnancy: In a well-nourished Canadian population the prevalence was 8%,[101] while it was 75% in rural India.[102] Despite the high prevalence of deficiency in low- and middle-income countries, B12-supplementation is not part of the World Health Organization's micronutrient supplementation recommendations.[103] Furthermore, the dosage of B12 included in the World Health Organization's multiple micronutrient supplement is arguably very low.[103,104] The principal nutritional source of B12 is found in animal products, such as meat and dairy products.[105] Because of this, women that consume little animal products are at increased risk of deficiency. Low consumption of animal products may be voluntary, or associated with cultural, religious or socioeconomic factors.[106] Rarer causes of B12-deficiency are related to malabsorption, such as Crohn's disease[107] and pernicious anemia.[108] B12 is actively transported across the placenta, with fetal levels of B12 higher than maternal B12 levels.[109,110] Levels of B12 are lower among

pregnant than non-pregnant women, which is only partly accounting for by the increased circulating blood volume.[111] Additionally, levels of B12 decline during pregnancy.[112]

Disregarding lack of evidence, B12-deficiency in pregnancy is often assumed to be associated with reduced fetal growth. There is especially one publication that is highly cited (70 times, June 2016, Web of Science), *"Low maternal vitamin B12 status is associated with intrauterine growth retardation in urban South Indians"* by Muthayya et al (2006).[100] There are multiple narrative reviews concluding that B12-deficiency is associated with reduced birth weight.[96,113,114] Yet, there are several other, less cited publications that report no association between B12 and birth weight.[115,116] The first systematic review on the topic was published while *Paper III* was under peer review.[7] However, that review was most likely biased (see *5.4.2 Systematic review*). As such, the literature on B12 in relation to birth weight and length of gestation may suffer from citation bias. Few studies report on the association between B12 and length of gestation. In *Paper III* we summarized and critically appraised the literature on the association between B12 and length of gestation and birth weight by conducting a systematic review. Given the high prevalence of B12-deficiency in pregnancy, and the potential ease of implementing flour fortification of B12[117] or antenatal B12-supplementation, research on this topic may have great public health interest.

1.6 Consequences of small birth size

The notion that nutrition and living conditions early in life may have life-long consequences was first forwarded by the Norwegian doctor Anders Forsdahl (1977), and later by Dr. David Barker (1986).[118,119] While Forsdahl focused on poverty in early childhood, Barker associated the fetal milieu, specifically related to maternal nutrition, with adult disease: "The association of ischaemic heart disease with neonatal mortality suggests that the childhood influences predisposing to it are related to nutrition during prenatal and early postnatal life."[119]

Barker hypothesized that unfavorable fetal environment, particularly poor maternal nutrition, would cause fetal "programming", later leading to chronic disease.[1] As an example of programming, he argued that the fetus would adapt to poor nutrition by inducing insulin resistance, which, in turn, in adult life is associated with DM and obesity.[77] In this framework, reduced birth size is an early symptom of a disease process. With the fetal insulin hypothesis, another framework was proposed:[120] A genetic variant could be associated with both fetal insulin resistance – causing restricted fetal growth – and insulin resistance in adulthood. Small size at birth was accordingly not a cause of later chronic disease, but an accompanying trait of an underlying condition. The general perception today is that both genes and fetal programming – now often referred to as *epigenetics* – can explain the association between small birth size and increased risk of adult chronic disease.[121]

The field of epigenetics, how cells use their genes, has received increasing attention in epidemiology. By altering the expression of genes, for instance by DNA methylation, environmental factors may cause disease later in life, and perhaps pass on traits to new generations.[122] Being an important methyl group donor, low levels of vitamin B12 in fetal life has been suggested to be associated with insulin resistance in childhood, mediated through epigenetic alterations. Epigenetic mechanisms have also been suggested in the link between maternal physical stress during pregnancy and brain development and later cognitive function.[123]

1.6.1 Brain development and cognitive function

The development of the fetal central nervous system is well coordinated and complex, but also sensitive to insults. The astrocytes, oligodendrocytes and microglia support the developing neurons in migration and modulation of synaptic connections.[124] Myelin, which enhances the transmission of action potentials along neurons, is produced by myelinating oligodendrocytes,

especially in late third trimester.[125,126] The precursor of the myelinating oligodendrocytes, the premyelinating oligodendrocytes, are especially vulnerable to toxins, nutritional deficiencies and hypoxia.[124] For instance, the methionine-homocysteine-pathway is suspected to be important for myelination, and maternal B12-deficiency in pregnancy has been associated with myelination disorders in their offspring.[127,128] Research suggest that in a situation of fetal hypoxemia and substrate deficiency, the brain is prioritized, known as the "brain sparing" effect.[68]

Preterm birth has been extensively associated with regional brain volume abnormalities in children.[129,130] Also children born SGA have been observed to have volume reductions in specific brain regions. In one study, cerebral cortical gray matter volume reduction was observed two weeks after birth among preterm SGA infants compared with newborns born preterm and weight appropriate-for-gestational-age (AGA).[131] At ages four to seven years, SGA children compared with AGA children were found to have reduced volumes of the cerebral and cerebellar gray and white matter, and volume reduction of the basal ganglia and smaller cortical surface area.[132] A study evaluating many of the same children studied in *Paper II*, however, found no regional cerebral volume reductions at 15 years among SGA children compared with controls.[130] The causative factors for insufficient brain maturation are still not settled, nor are the predisposing factors for later cognitive function.

There are conflicting results regarding the causes of reduced cognitive function in childhood: Both genetic and environmental factors are important.[133] LBW has in itself been associated with reduced intelligence quotient (IQ) in childhood, regardless of maternal smoking habits for instance.[134] Preterm birth is associated with reduced cognitive abilities in childhood, adolescence and adult life.[135–137] SGA children, compared with controls, have been observed to have lower IQ scores at 10 years, although of uncertain clinical significance.[138] Similar findings have been observed at 17 and 26 years.[139,140] Another study, though, found no differences in IQ

score between SGA and AGA children at 7 years:[141] This study, however, applied a less conservative definition for SGA than the other studies. The varying definitions of SGA illustrate one of the problems when comparing epidemiological studies of fetal growth and birth weight. Disparate study populations and outcome measures further add to the uncertainty.

How fetal growth is biologically related to later cognitive function is unclear. Reduced volumes of the hippocampus, the cerebellum and the thalamus, among other brain structures, have been associated with cognitive deficits.[142–145] Other studies have failed to find any correlations between IQ and brain measures.[146]

2. Aims

The main aims of this thesis were to explore whether maternal blood concentrations of two important micronutrients, glucose and vitamin B12, were associated with birth size, and evaluate how birth size is associated with later neurological and cognitive development. The following aims were investigated to help elucidate the overall aims:

- To investigate how maternal glucose tolerance in late pregnancy is associated with birth size, with special emphasis on birth size in women with high tolerance.
- To systematically review and re-analyze the published literature on studies reporting on the association between maternal vitamin B12 in pregnancy and birth weight and length of gestation.
- To study whether small-for-gestational-age (SGA) children have reduced cognitive function and regional brain volume reductions in childhood and adolescence, and explore potential associations between regional brain volumes and cognitive function. Also, to evaluate whether cognitive scores and regional brain volumes in the SGA group varied between those with fetal growth restriction (FGR) and those without FGR.

3. Material and methods

This thesis is based on two distinct studies and methodological approaches. *Paper I* and *II* are traditional longitudinal cohort studies, based on the same original study. The final paper, *Paper III*, is a systematic review with individual patient data (IPD) meta-analysis, and is based on multiple longitudinal cohort studies. For readability, the *Paper I* and *II*, and *Paper III,* will therefore be discussed under separate subheadings in each section.

3.1 Study design

3.1.1 Scandinavian SGA-study (Paper I and II)

The two first publications in this thesis were based on the *Scandinavian Small-for-Gestational-Age (SGA) Study* (Figure 3.1). It was a large, prospective, multicenter study conducted in Trondheim and Bergen, Norway, and Uppsala, Sweden, between January 1986 and March 1988.[55] The National Institute of Child Health and Human Development funded the study. The overall aims of the *Scandinavian SGA-study* were to study fetal growth, perinatal and later outcome, and the tendency to repeat a negative outcome in consecutive births.

Referral of eligible patients was done by general practitioners and obstetricians in Trondheim and Bergen, and by antenatal care centers in Uppsala. Clinical and follow-up data were collected at the university hospitals. The women were screened in gestational weeks 17, 25, 33 and 37, and at delivery. There was a follow-up study of many of the offspring up to five years of age, and among children from Trondheim up to 15 years of age (see *3.2.2 Paper II*).

Figure 3.1 Illustration from recruitment brochure of the *Scandinavian SGA-study*.

3.1.2 Systematic review (Paper III)

As detailed in section *1.3.2 Systematic reviews*, the hallmarks of a systematic review are a clearly formulated question, systematic methods are used to identify, select and critically appraise relevant research, and that data from relevant research is collected and analyzed.[26] The process should be explicit and reproducible. Due to high level of reporting bias along with potential sources of confounding, we deemed it necessary to perform an IPD meta-analysis. The protocol for the systematic review is available at Prospero.[147] The manuscript under consideration for publication was reported according to the PRISMA guidelines.[44]

The systematic review was originally a collaborative study between the Norwegian University of Science and Technology (NTNU) and Yale University (Figure 3.2). The study was initiated in September 2013, and a manuscript for a systematic review not including a meta-analysis was completed January 2014. Because of great variation in reporting of results – if reported at all –quantitative assessment of the literature was deemed inappropriate. To properly address the research question, the scope of the systematic review was expanded, and we decided to collect the necessary data from the individual studies (see below). Due to the need for close collaboration with research groups that provided data from individual studies, the review group grew in size. At submission, the review was a collaborative study between 14 research groups from eight countries. The research group at the Generation R Study at Erasmus MC, Rotterdam, the Netherlands, deserves a special mention, as it became one of the key contributors and hosted multiple visits during the later stages of the study.

Figure 3.2 Main contributors to the systematic review on the association between vitamin B12 in pregnancy and risk of preterm birth and low birth weight. MC, medical center; NTNU, Norwegian University of Science and Technology.

The topic of interest was how vitamin B12 (B12) in pregnancy affects length of gestation and birth weight. Because no randomized controlled trials on vitamin B12 supplementation in pregnancy were identified at the beginning of this project, a review of observational studies was necessary. The final research question was as follows: "Are maternal blood levels of vitamin B12 in pregnancy associated with length of gestation and newborn birth weight?".

Inclusion criteria

Only studies on humans were considered. In order to reduce bias and to be able to estimate risk ratios (RR) – rather than odds ratios – only longitudinal cohort studies were eligible. B12 had to be measured in maternal serum or plasma after conception and before the day of delivery. Birth weight had to be measured at birth. To represent a more general population, studies were excluded if mainly women or offspring with specific conditions, such as preeclampsia or congenital malformations, were evaluated. Because close collaboration with researchers providing data from individual studies was important, we only included studies published in 1998 or later. Studies with fewer than 50 pregnancies would have limited value in the analysis, and were not considered for the review.

Search methods

The service providers used for the electronic search were PubMed, Scopus, Web of Knowledge, EBSCO and OvidSP. The latter was used to access the databases of MEDLINE, EMBASE and GLOBAL HEALTH. CINHAL was accessed by use of EBSCO. A combination of text words and controlled vocabulary (e.g. MESH) was used. Individual search strings were adapted to each service provider and database, and consisted of the following four elements: *B12* AND *pregnancy* AND *birth weight/length of gestation* NOT *restriction terms*. Restriction terms were added to limit the number of irrelevant publications, and were constructed for: review articles, intervention studies and case reports, and studies evaluating old adults, children (other than infants), rodents and patients with anemia. Notably, no language restrictions were applied. The exact search strings for each service provider and database may be found in the Appendix. Presented below is the exact search term used for PubMed:

("Vitamin B 12"[Mesh] OR B12[Text Word] OR "B 12"[Text Word] OR cobalamin[Text Word]) AND (pregnan*[Text Word] OR Pregnancy[Mesh] OR gestation*[Text Word] OR fetus[MeSH] OR fetus*[Text Word] OR foetus*[Text Word] OR foetal* [Text Word] OR fetal*[Text Word] OR "Fetal Development"[Mesh] OR "Infant, Newborn"[Mesh]) AND ("Infant, Low Birth Weight"[Mesh] OR "Birth Weight"[Mesh] OR birthweight[Text Word] OR "birth weight"[Text Word] OR SGA[Text Word] OR "fetal growth retardation"[MeSH] OR IUGR[Text Word] OR "growth restriction"[Text Word] OR "growth retardation"[Text Word] OR "small for gestational age"[Text Word] OR "small for date"[Text Word] OR "Infant, Premature"[Mesh] OR "Premature Birth"[Mesh] OR "Gestational Age"[Mesh] OR preterm[Text Word] OR prematur*[Text Word] OR "gestational age"[Text Word] OR "length of gestation"[Text Word] OR "duration of pregnancy"[Text Word]) NOT ("Review"[Publication Type] OR "Child"[Mesh] OR "Aged"[Mesh] OR "Case Reports"[Publication Type] OR "Clinical Trial"[Publication Type] OR "Rodentia"[Mesh] OR "Anemia"[Mesh])*

The electronic search was last conducted in August 2015. In addition to the electronic search, the reference lists of all studies read in full text were carefully examined to find eligible reports not otherwise identified.

Selection process of studies

Adapted from the Cochrane Handbook for Systematic Reviews of Interventions,[148] the following selection process was followed:

1) *Electronic and gray literature searches.* The electronic search was carried out by the first author (Tormod Rogne (TR)), and gray literature searches were carried out by two authors (TR, and Kari R Risnes (KRR) or Myrte J Tielemans (MJT)).
2) *Examination of titles and abstracts to remove irrelevant reports.* Duplicates from the electronic searches were removed in the reference manager. With an over-inclusive approach, eligibility of all references was thereafter assessed by screening of titles and abstracts by the first author (TR).
3) *Examination of full text of remaining studies.* All potentially eligible studies were read in full text and assessed independently by two review authors (TR, and KRR or MJT). The publication was excluded if it met one of the following exclusion criteria: animal study; not a prospective cohort study; B12 not measured, or in most cases measured before conception or at or after delivery; B12 not measured in mother's blood; no information on birth weight or length of gestation; study population constituted by mothers or offspring with specific disorders (other than B12-deficiency) or genetic variants; a newer or more complete report on the same study was included; no information on the association between maternal blood levels of B12 in pregnancy and birth weight or length of gestation presented in the published reports or provided by the authors; published prior to 1998.

At least two authors from each of the eligible published reports were contacted in order IPD. All authors were contacted at least three times. IPD was received without personal identification. If IPD was not available, the authors were asked to re-analyze their data and provide results from requested analyses. When neither IPD nor re-analyses could be retrieved, relevant estimates were extracted from the publications.

Quality assessment

The risk of bias of included studies was assessed by use of the Newcastle-Ottawa scale, which was modified by TR, KRR and MJT to be suited for the systematic review.[149] The checklist contains seven questions, where each item gives one or zero points. The scale goes from zero to seven, with the latter designating the lowest risk of bias. A study was considered to have low risk of bias if it scored four or more, and high risk of bias otherwise. Two review authors did the risk of bias assessment (TR and MJT). Disagreements were resolved by consulting a third author (KRR). The complete scale may be found in the Appendix.

3.2 Study population

Paper I considered the first part of the *Scandinavian SGA-study*, that is, from early pregnancy through delivery. As follows, fetal growth and newborn birth weight were the *outcome* measures of interest. In the ensuing *Paper II*, the follow-up studies were considered. Accordingly, fetal growth and newborn birth weight were *exposures* in this paper. Importantly, although there was an overlap in the study populations evaluated in *Paper I* and *Paper II*, the latter publication was independent from the first, that is, glucose tolerance was not considered in *Paper II*. *Paper III* evaluated study populations from multiple individual studies. The flow of studies included in the systematic review will briefly be covered in the results section (*4.3 Paper III*).

In the *Scandinavian SGA-study*, women were eligible if they were of Caucasian origin, spoke one of the Scandinavian languages, expected their second or third child, and had a singleton pregnancy.

A total of 6,354 women were recruited to the study at the first antenatal visit (Figure 3.3). Three groups were constructed based on the eligible women who made their first appointment prior to gestational week 20 (n = 5,722): (1) a 10% random sample, constructed

to serve as a population reference, (2) a low risk group, and (3) a high risk group. The random sample was constructed using the sealed envelope method. Women were categorized as high risk if they fulfilled at least one of the following criteria: (1) previous low birth weight (LBW) child, (2) previous perinatal death, (3) maternal low pre-pregnancy weight (<50 kg), (4) smoking around the time of conception, and (5) chronic maternal hypertension or renal disease. The low risk group served as a "rest population" and consisted of women that did not have any of the mentioned risk criteria, and were not randomly selected to the random sample. All women in the high risk group and the random sample were invited to detailed follow-up during pregnancy.

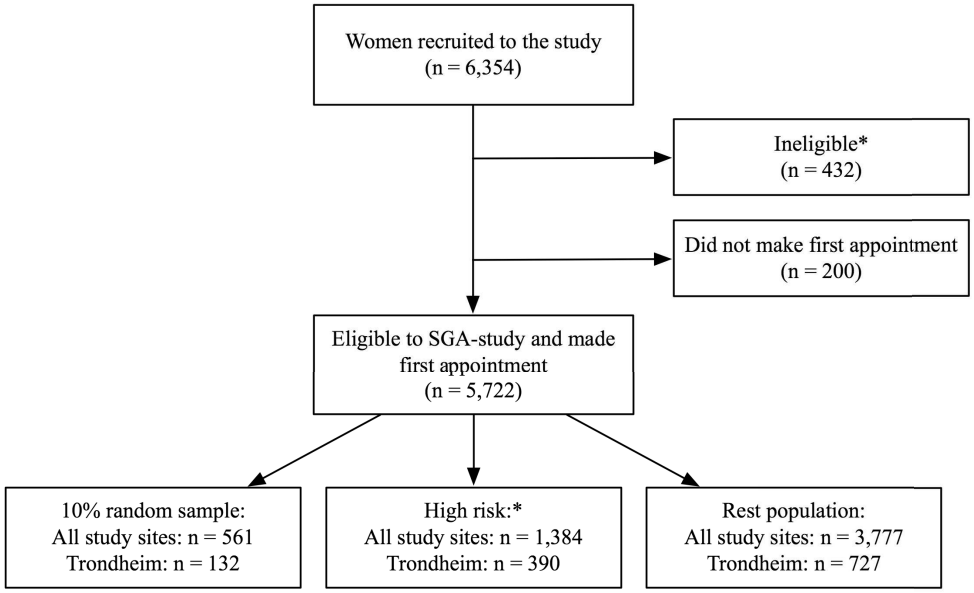


Figure 3.3 Flow chart illustrating the construction of the three study groups in the *Scandinavian SGA-study*. *See text in section *3.2 Study population*.

3.2.1 Paper I

A total of 1,945 women were eligible for detailed follow-up, of which 860 (44%) completed an oral glucose tolerance test (OGTT) in week 37 according to protocol, and regardless of risk factors for gestational diabetes mellitus (GDM; Figure 3.4). Women with pre-pregnancy overt

diabetes mellitus (DM) were excluded (n=3), as were pregnancies where an OGTT was not carried out (n=1,085). Only two of the remaining pregnancies delivered preterm (birth before 37 completed weeks), and these were excluded to avoid mixing of effects of preterm birth and restricted fetal growth. The final study population consisted of 855 pregnancies. Of these, 30% were derived from the random sample, while the remaining 70% had an increased risk for SGA birth.

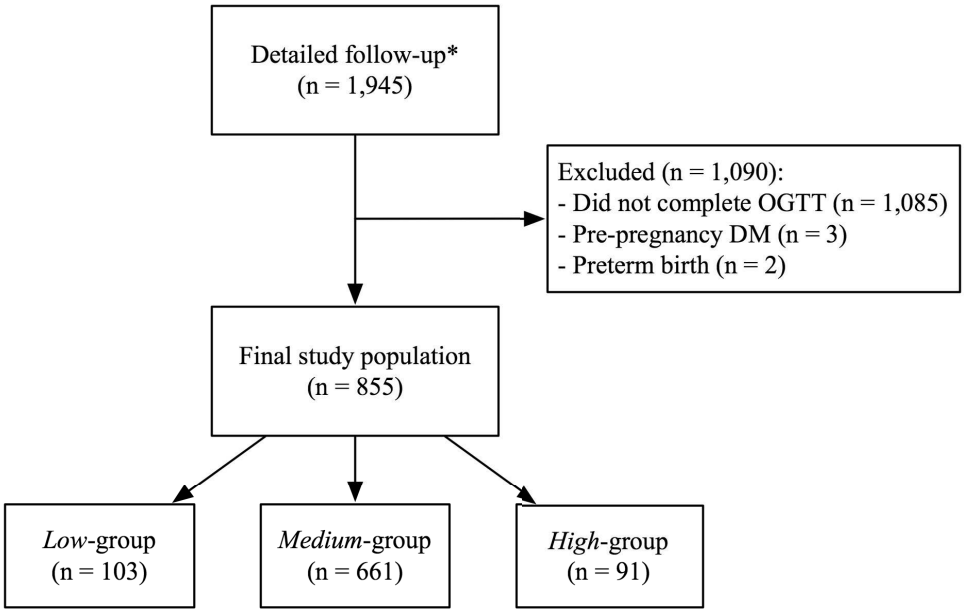


Figure 3.4 Flow chart illustrating the allocation of participants in *Paper I*. *Eligibility criteria described in *3.2 Study population*. DM, diabetes mellitus; OGTT, oral glucose tolerance test.

Finally, the population was divided into three study groups based on blood glucose increase after the OGTT (see *3.3.2 Glucose*): The 10% of women with the lowest increase (the *Low*-group (n=103)), the 10% of women with the highest increase (the *High*-group (n=91)), and the group of women with moderate increase (the *Medium*-group (n=661)). Focus was on women with the lowest increase.

3.2.2 Paper II

All children born SGA (including children born SGA in the rest population) were invited to a follow-up study, the *Scandinavian SGA II-study*, at one and five years at the three study sites.[55] Additionally, any child from the random sample were invited to follow-up. At nine years, children born SGA or from the random sample were invited to a separate follow-up study in Trondheim.[150] Yet another study invited the same children from Trondheim to a follow-up study at 15 years.[151] *Paper II* did not consider the examinations done at one year of age. As findings at 9 and 15 years were of particular interest in, only children from Trondheim were considered in *Paper II* (Figure 3.5).

To avoid mixing of effects of preterm birth and small birth size, and as the definition of fetal growth restriction (FGR) relied on ultrasound measurements in gestational week 37, preterm deliveries (birth before 37 completed weeks) were excluded, as were pregnancies with missing information on length of gestation. Children with congenital malformations or cerebral palsy (due to severe birth asphyxia) were also excluded. Twenty-nine of the eligible children did not participate in the follow-up study.

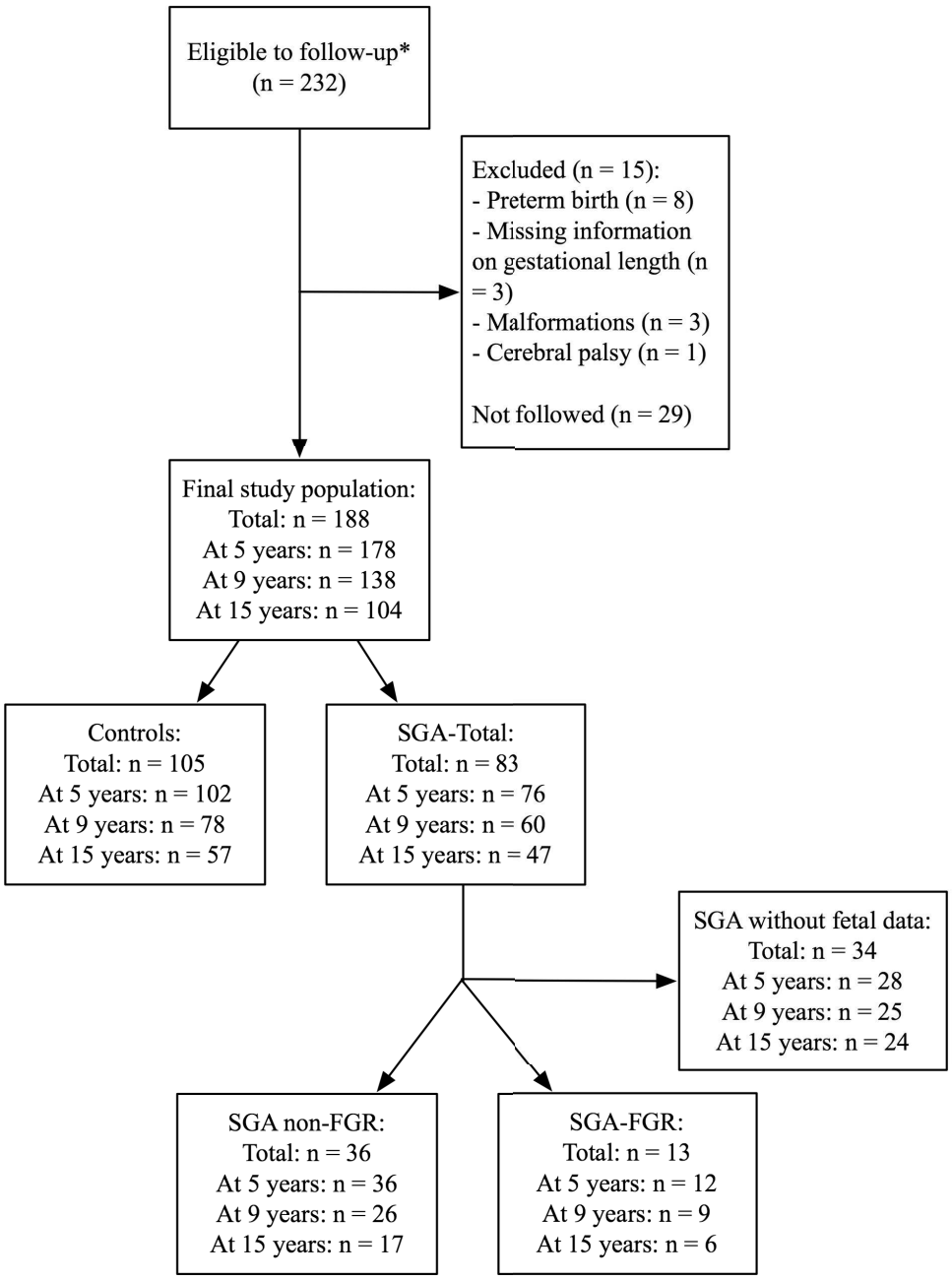


Figure 3.5 Flow chart illustrating the allocation of participants in *Paper II*. *Random sample or born SGA. FGR, fetal growth restriction; SGA, small-for-gestational-age.

The children were first divided into two study groups: an SGA group (SGA-Total; n=83) and a non-SGA group (Controls; n=105) serving as controls (see *3.3.1 Fetal growth*). The children born SGA were further divided into an SGA group with restricted fetal growth, (SGA-FGR; n = 13), and an SGA group without restricted fetal growth (SGA non-FGR; n = 36), based on serial ultrasound measurements. Many children had no fetal growth data (n = 34); this was especially true for children derived from the rest population who, by design, were not followed during pregnancy. Figure 3.5 reports the number of subjects at each visit.

3.3 Clinical variables

The clinical variables of the *Scandinavian SGA-study* (*Paper I* and *II*) will be discussed under each of the subheadings of this section. The last subheading (*3.3.7 Variables included in systematic review*) covers the variables in the systematic review (*Paper III*).

3.3.1 Fetal growth

Ultrasound examinations were performed in weeks 17, 25, 33 and 37 of gestation by specially trained and highly qualified midwives.

eSnurra is a well-established mathematical model developed by the National Center for Fetal Medicine, St. Olav's University Hospital and Norwegian University of Science and Technology (NTNU), Trondheim.[152] We used it to estimate gestational age and expected date of delivery from ultrasound measurements of biparietal diameter (BPD) at 17 weeks. Based on BPD and mean abdominal diameter, eSnurra was also used to estimate fetal weights at weeks 25, 33 and 37. Femur length was used when information on BPD was missing (n=19 in *Paper I*, none in *Paper II*). Fetal weight prior to day 167 was manually estimated by use of the *eSnurra plastic wheel,* as the eSnurra mathematical model does not estimate weights this early in

pregnancy (n=14 in *Paper I*, n=2 in *Paper II*).[153] A *percent* deviation for weight was also estimated from a population based reference at each gestational age.[152]

Per protocol, SGA was defined as birth weight below the 10th percentile for a given gestational age, accounting for parity and sex. FGR was in *Paper I* defined as growth below the 10th percentile between week 25 and 37, based on the growth in the random sample (i.e. the "population reference" of the *Scandinavian SGA-study*). In *Paper II*, FGR was defined as growth from week 25 to week 37 more than two standard deviations (SD) below the mean of the non-SGA group (i.e. the "control group" in *Paper II*).

3.3.2 Glucose

The exposure of interest in *Paper I* was the maternal response to an oral glucose challenge. After an overnight fast, the OGTT was administered in gestational week 37. The validity of an OGTT this late in pregnancy has been demonstrated to be very high.[154] Capillary glucose values were recorded after an overnight fast, and two hours after administering 75 g glucose in 200 mL water (consumed over 10 minutes). We were mainly interested in women at the opposite end of the glucose tolerance spectrum from those with GDM or DM. We therefore identified glucose tolerant pregnant women, meaning that the change in blood glucose from fasting state to two hours after the glucose administration is relatively low. We defined delta (Δ) glucose (mmol/L) as two-hour glucose *minus* fasting glucose. A large Δ glucose has previously been linked with deterioration of beta-cell function, while high insulin sensitivity and glucose tolerance have been associated with a low Δ glucose.[94] Except for rare cases of postprandial hypoglycemia, there are no generally accepted cut-offs for a low Δ glucose. For this reason, we categorized our population based on the normal distribution: From the 10% random sample (i.e. the population reference), we derived cut-offs for the 10th and the 90th percentiles of the Δ glucose distribution, which were 0.8 mmol/L and 3.6 mmol/L, respectively. Thus, three groups

were identified according to the change in blood glucose: *Low* (*L*; Δ glucose ≤ 0.8 mmol/L (i.e. <10th percentile); n = 103; 12%); *Medium* (*M*; 0.8 mmol/L < Δ glucose < 3.6 mmol/L (i.e. >10th percentile and <90th percentile); n = 661; 77%); and *High* (*H*; Δ glucose ≥ 3.6 mmol/L (i.e. >90th percentile); n = 91; 11%).

To evaluate the validity of our statistically defined glucose tolerance groups, for each group, we tabulated the frequencies of the more well recognized glucose metabolism disorders presented in Table 3.1.[6,155]

Table 3.1 Glucose metabolism disorders (adapted to capillary values).

	Fasting (mmol/L)	Two-hour (mmol/L)
Hypoglycemia	≤2.80	≤2.80
Impaired fasting glycemia	≥5.6 and <6.1	<7.8
GDM	<6.1	≥7.8 and <11.1
DM	≥6.1	≥11.1

DM – overt diabetes mellitus; GDM – gestational diabetes mellitus.

3.3.3 Maternal clinical variables

The maternal clinical variables used in *Paper I* and *II* were prospectively collected during pregnancy. At the first study visit, the mother's social, medical and family history, dietary information and smoking habits were recorded, maternal height and weight were measured, and blood samples were collected.[55] For the women included in the detailed follow-up, further blood samples and habits and events during pregnancy were collected at study visits in week 25, 33 and 37. Because the adverse effect of smoking on fetal growth may be greatest in late pregnancy,[156] we classified smokers based on reported smoking habits in week 33. Smoking habits were categorized into 0, 1-9, 10-19 and ≥20 cigarettes a day. Weight gain was recorded as the change in weight from the one reported prior to pregnancy, to the one at the time of the examination of interest. Highest level of completed education was stratified into primary school (≤ 9 years), middle school (10-12 years) and high school (≥ 13 years).

3.3.4 Offspring clinical variables

The following offspring characteristics of interest were recorded at birth: Sex, birth weight and length, head circumference, triceps skinfold (measure of subcutaneous fat), ponderal index (100*birth weight (g)/crown-heel-length^3 (cm)), placental weight, and length of gestation (see *3.3.1 Fetal growth)*. The physical condition of the newborn was assessed by the 10-point Apgar score one and five minutes after delivery.

3.3.5 Cognitive function

Cognitive function was of interest in *Paper II*, and was assessed by the Wechsler Preschool and Primary Scale of Intelligence Revised (WPPSI-R) at five years,[157] and the Wechsler Intelligence Scale for Children Revised (WISC-R) at 9 years.[158] The same, specially trained psychometrist examined all children, and was blind to SGA status.

Both the WPPSI-R and the WISC-R yield performance, verbal and full intelligence quotient (IQ) scores. Performance IQ is associated with abstract problem solving, while verbal comprehension and working memory are important to verbal IQ. Both performance IQ and verbal IQ were composite scores of multiple subtests. Full IQ was a combination of verbal IQ and performance IQ. The estimated scores were age specific.

3.3.6 Magnetic resonance imaging

In *Paper II*, brain volumes were measured at age 15 by use of magnetic resonance imaging (MRI). Details of the MRI technique has been detailed in a previous paper, and will be summarized here:[159] The MRI machine was a 1.5-Tesla Siemens Symphony Sonata (Siemens AG, Erlangen, Germany). For the morphometric analysis the imaging was a 3D inversion recovery prepared at a fast low flip angle gradient echo sequence (MP-RAGE) with 128 sagittal partitions, 1.33 mm slice thickness, TR between inversion pulses of 2730 ms, TR/TE/flip

angle/TI: 7.1 ms/3.45 ms/7°/1000 ms, an acquisition matric of 256 x 192 x 128, square FOV of 256 mm, NEX 1, and an acquisition duration of 8.5 minutes.

A technique developed by Fischl et al. was applied for the automated labeling of human brain structures.[160] We obtained volumes for total brain and cerebral and cerebellar white matter and a series of gray matter structures including the thalamus, the hippocampus, the amygdala, and the cerebral and cerebellar cortices.

All brain volumes were normalized by total intracranial volume by use of the covariance method described by Jack et al.:[161]

NV = OV – RLG(TIV – TIVmean),

where NV was the normalized volume of the structure of interest, OV the original volume, RLG the regression line gradient between the original volume and the total intracranial volume, TIV the subject's total intracranial volumes, and TIVmean the total intracranial volume in the control group.

3.3.7 Variables included in systematic review

By design, all clinical variables were collected by the individual studies included in the review. As a consequence, one may assume a certain level of inter-study variation, e.g. use of different scales to weigh the mother. In the collection of data from the individual studies we applied a pragmatic approach and asked for an appropriate, but limited number of variables. The level of detailed information varied. The full list of variables asked for is described in the protocol.[147] Below are listed the exposures, outcomes, and covariates that constituted the clinical variables that were eventually used in the review.

Exposure

The exposure of interest was maternal vitamin B12 in serum or plasma during pregnancy. Indirect measures of B12 levels, e.g. methylmalonic acid and homocysteine, were not considered in this study. B12 was measured by use of radioimmunoassay, electroluminescence assay, or microbiologic assay. To standardize B12 across the length of pregnancy, we calculated B12 SD scores for each trimester based on studies providing IPD and re-analyzed aggregate data. B12-deficiency was pre-defined as <148 pmol/L (<200 pg/L).[162] For the sensitivity analyses, we constructed B12 tertiles based on IPD: <148 pmol/L (tertile 1), 148-216 pmol/L (tertile 2), and >216 pmol/L (tertile 3).

Outcomes

The primary outcomes of interest were birth weight in grams (continuous), LBW (birth weight <2,500 g), and SGA (defined below).[10] Weight had to be measured at birth and could not be based on retrospective reported data. Four secondary outcomes were defined *a priori*: length of gestation in days (continuous), preterm birth (birth before 37 completed weeks), very preterm birth (birth before 32 completed weeks), and very low birth weight (birth weight <1,500 g). Length of gestation had to be recorded at birth and needed to be estimated by use of ultrasound measurements or last menstrual period, or a combination of the two. An additional outcome was later added; birth weight SD score. This outcome was calculated by use of the INTERGROWTH 21st Project, which estimates an SD score based on gestational age and sex specific birth weight reference standards.[163] SGA was defined as birth weight SD score below the 10th centile. We assumed birth weight SD score to serve as a proxy of fetal growth. Likewise, SGA was assumed a proxy of restricted fetal growth.

Covariates

The following continuous covariates were considered: maternal age and pre-pregnancy or pregnancy body mass index (BMI) and weight. We also considered the dichotomous outcomes parity (nulliparous vs. primiparous or multiparous), smoking habits (smoking during pregnancy vs. not smoking during pregnancy), and highest completed education (completed high school, equal to 13 years of education, vs. not completed high school). To be included, maternal weight had to be measured as part of the original study, and not by self-report. BMI was calculated using the formula: weight (kg)/height $(m)^2$. BMI was preferred over weight. The earliest measurement of maternal anthropometry was used, preferably pre-pregnancy. Information on parity, smoking habits and education were in the individual studies collected by use of questionnaires.

3.4 Statistical analysis

All statistical analyses were carried out using Stata SE version 12.1 (*Paper I* and *II*) and version 13.1 (*Paper III*), Stata Corporation, College Station, TX, USA. Precision was assessed by 95% confidence intervals (CI), and statistical significance was defined as $p<0.05$. In *Paper II*, p-values were corrected for multiple comparison by use of the Benjamini-Yekutieli method in the multivariable analyses of the subgroup analyses of brain volumes and in the pairwise correlation analyses.[164]

3.4.1 Univariable analysis

In all papers in this thesis, normally distributed variables were presented as means, with 95% CI. Non-normally distributed variables were presented as medians, with corresponding 25^{th}- to 75^{th}-percentiles, or range. Frequencies were presented as the number of observations and its proportion of total.

In *Paper I,* generalized linear models were applied in both the analysis of continuous as well as binary outcome variables. Linear regression models were fitted in the former situation, and Poisson regression models in the latter. Some of the continuous outcome variables in *Paper I* were not normally distributed, and to avoid violation of the linearity assumption, we carried out these analyses by use of appropriate non-identity link functions. For the sake of consistency, we carried out all analyses of continuous data in *Paper I* with a generalized linear model. To relax the assumption of distribution and independence of errors, we used regression with robust standard errors. We expected that the binary outcomes under analysis in *Paper I* had an incidence of more than 10%, which made logistic regression unsuitable.[165,166] We instead applied a Poisson regression model with robust error variance. It has been argued that this is a more appropriate method for the analysis of binary outcomes in prospective studies.[166,167]

In *Paper II*, group means were compared by independent *t* tests, medians by Wilcoxon-Mann-Whitney tests, and frequencies by chi-squared test or Fisher's exact test (the latter if an expected frequency in any cell of five or less). Association between IQ scores and brain volumes were explored by the use of pairwise correlation analyses. After transforming non-normally distributed outcome variables to an appropriate scale, continuous outcome variables in *Paper II* and *III* were analyzed by use of regular linear regression. Some of the individual studies included in *Paper III* had high incidences of LBW, SGA and preterm birth. Binary outcomes in *Paper III* were therefore analyzed by the same technique as in *Paper I*, that is, by use of Poisson regression with robust error variance.

3.4.2 Multivariable analysis

In order to adjust for potential confounding, multivariable analyses were utilized in all papers. All models were tested for normality of residuals, heteroscedasticity and multicollinearity. Non-

normally distributed independent variables were transformed to an appropriate scale prior to analysis.

Maternal smoking habits during pregnancy was identified as a potential confounding factor for the association between glucose tolerance and fetal growth (*Paper I*), as smoking may affect both glucose tolerance[168] and birth weight (Figure 3.6).[169] In *Paper II*, we did not treat smoking as a confounder in the analyses of fetal growth on cognitive function and brain volumes.[170] Some of the detrimental effects of smoking on brain development and cognitive function may be mediated through fetal growth. Indeed, researchers have argued that maternal smoking during pregnancy has no direct causal effect on IQ of the offspring.[134]

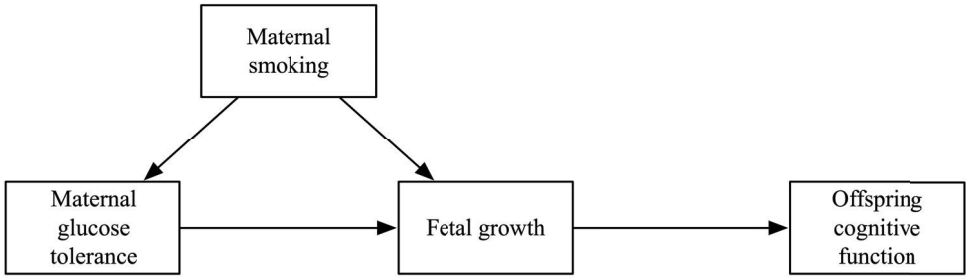


Figure 3.6 The suspected causal associations between maternal glucose tolerance, fetal growth, offspring cognitive function and smoking habits.

The following potential confounders were considered in the multivariable analyses of *Paper I:* Previous LBW child, mother's highest completed education, maternal age, parity, pre-pregnancy BMI, weight gain, smoking during pregnancy, gestational age when the OGTT was performed, gestational age at birth, and sex of the baby.

Three potential confounders were identified *a priori* in *Paper II:* Mother's highest completed education, offspring sex and age at examination. The two latter were adjusted for in the multivariable analysis of brain volumes, while in the analysis of cognitive function, maternal education and child sex were taken into account (IQ scores were age specific).

In *Paper III* we selected the potential confounders based on a combination of clinical and epidemiological knowledge of factors associated with exposure and outcome. While a model adjusting for maternal pre-pregnancy BMI, age, parity, education and smoking habits was preferred, this would leave us with only a few studies under analysis. We decided that the best multivariable model based on the data at hand was one including maternal age, BMI and parity. The earliest measurement of BMI was preferred, and when BMI was unavailable, we used maternal weight instead. Alternative multivariable models were explored in sensitivity analyses (see below).

We did not use interaction terms in the analyses of any of the papers.

3.4.3 Meta-analysis

On study-level, as described above, mean difference of the continuous outcomes birth weight (g), gestational age at delivery (days) and birth weight SD score were analyzed by linear regression. RR for the outcomes LBW, SGA and preterm birth were estimated by Poisson regression with robust error variance. Results from the individual studies were pooled in an IPD meta-analysis. IPD meta-analyses can either be carried out in a "one-step" or "two-step" approach:[171] In a two-step approach, summary statistics are generated at study level (step one), and these estimates are then pooled using traditional meta-analytical approaches (step two). This is the most widely used approach.[171] Alternatively, one may pool all the raw data into one model, using a one-step approach, and in the regression model take clustering by study into account. This approach is more difficult, but some argue that it is a more exact statistical approach.[172] Others argue that there is little difference in results between the two techniques.[171] The data in *Paper III* was analyzed using a two-step approach. Aggregate data from individual studies were included in the analyses when IPD was not available. The studies contributing in the analyses were weighed using the DerSimonian & Laird random-effects model, assuming

that in addition to within-study variance, there was also between-study variance (as opposed to a fixed effect model assuming only within-study variance).[173]

Both univariable and multivariable models were built. In the main multivariable model we adjusted for maternal age, BMI (or weight when BMI was unavailable) and parity. When IPD was unavailable, the following re-analyses were requested from the individual studies: linear regression of B12 and B12-deficiency in relation with birth weight and gestational age at delivery; and Poisson regression of B12 and B12-deficiency in relation with LBW, SGA and preterm birth. Both crude and two multivariable models were requested (adjusting for maternal age, BMI and parity; and adjusting for maternal age, BMI, parity and smoking habits). Relevant results were extracted from the publications when neither IPD nor results from requested re-analyses were available.

We explored subgroup effects and heterogeneity by conducting multiple sensitivity analyses. Among the sensitivity analyses defined per protocol, the impact of individual studies' risk of bias was explored by stratifying the analyses into high risk versus low risk of bias. Demographic differences were explored by subgroup analysis of country income category; high-income country vs. low- and middle-income countries, as defined by The World Bank.[174] Because the maternal level of B12 declines throughout pregnancy,[112] and there are reports of varying association between B12 and birth weight depending on time of B12 ascertainment,[100] we stratified the analyses by 1st, 2nd and 3rd trimesters, and 1st and 2nd trimesters combined. A final *a priori* analysis was conducted excluding each of the studies one by one, to explore the impact individual studies had on effect estimates and heterogeneity. Additionally, we carried out several post hoc sensitivity analyses. Lower B12 in pregnancy has been associated with maternal obesity,[175,176] hence we stratified our analyses into overweight (BMI $\geq$25 kg/m^2) and non-overweight (BMI <25 kg/m^2). Different methods of measuring B12 may yield slightly different results, which is why we performed subgroup analyses of B12 assay technique

(radioimmunoassay, electroluminescence, microbiological).[177] We also explored how choice of statistical method affected the analyses, with sensitivity analyses of alternative multivariable models (e.g. more saturated model including maternal education and smoking habits in addition to the main model), fixed effects model, Poisson regression with non-robust error variance, and logistic regression model. Lastly, we evaluated the impact of excluding studies that only evaluated term births. Inspired by a recent publication,[175] we performed a post hoc meta-analysis of the association between B12 and maternal weight.

A funnel plot is a scatter plot of the effect size of an association from individual studies, against the precision of the studies.[178] Larger studies tend to be closer to the true effect size, while smaller studies tend to scatter more. If there are more studies on one side of the "true effect" line, one may suspect presence of publication bias. We evaluated presence of publication bias by use of this method. The degree of inconsistency, i.e. heterogeneity, across studies in the meta-analyses was described using the I^2-statistic.[179] We considered heterogeneity to be present when I^2 was greater than 30%, meaning that more than 30% of the total variation across studies was due to heterogeneity.

3.4.4 Multiple imputation

Because of a relatively large amount of missing data of the independent variables in *Paper I*, we applied multiple imputation.[180,181] We followed recommended guidelines for the reporting of multiple imputation analysis.[180]

We assumed that the missing data was due to missing at random. In the imputation model we included all variables included in any of the regression analyses or descriptive tables, and 21 auxiliary variables (including defined risk factors for SGA birth). No categorical variables with more than two levels were included in the imputation model, no variables were rounded, and there were no interaction terms.[182] The native Stata command *mi impute* was used

to impute 20 datasets. Necessary transformations of non-normally distributed independent variables were executed after imputations. Results from the complete case analyses were reported in supplementary material for comparison.

Although there was considerable missing information on *Paper II* as well, especially maternal education and fetal growth, we know that this information was not missing at random. Indeed, missing information was almost exclusively restricted to the group of women who were not followed during pregnancy, i.e. the women in the low risk population. Because of this, multiple imputation was not appropriate in *Paper II*.

There are discrepant opinions as to whether imputation is appropriate in the setting of meta-analysis of IPD.[183,184] We decided to refrain from doing multiple imputation in *Paper III*.

3.5 Ethics

All participants in the *Scandinavian SGA-study* were informed about the study aims and objectives, and gave their written consent. The Regional Committee for Medical and Health Research Ethics approved the *Scandinavian SGA-study* (references JSV/BLF 19.03.1984; 86/91/TK; and 4.2005.2605) and the systematic review (reference 2014/615/REK midt). The individual studies included in the systematic review were approved by their respective regional ethics committees.

4. Results

4.1 Paper I

Rogne T, Jacobsen GW. Association between low blood glucose increase during glucose tolerance tests in pregnancy and impaired fetal growth. Acta Obstet Gynecol Scand; 2014;93:1160-9.

The constructed glucose tolerance groups reflected clinically familiar glucose metabolism disorders: There were few cases (5%) of gestational diabetes mellitus or overt diabetes mellitus in the group of women with the lowest glucose increase (the *Low*-group), but these conditions were present in 96% of pregnancies in the group of women with the highest glucose increase (the *High*-group).

The fetal growth in the *Low*-group deviated increasingly more in a negative direction from gestational week 25 to 37 than both the *High*-group and the *Medium*-group. The group mean differences of deviation from expected fetal weight were greatest in gestational week 37: *Low* vs. *Medium* -2.4 percentage points (95% confidence interval (CI) -4.0, -0.8); *High* vs. *Medium* 2.3 percentage points (95% CI 0.4, 4.2). Newborns in the *Low*-group were thinner than those in the *Medium*-group, as judged by the ponderal index (-0.1 g*100/cm^3 (95% CI -0.1, 0.0)) and the triceps skin fold thickness (-0.3 mm (95% CI -0.4, -0.1)). However, there were no differences in birth weight or risk of small-for-gestational-age (SGA) birth between the groups, but a strong tendency of decreasing birth weight and risk of SGA birth with increasing glucose intolerance.

4.2 Paper II

Rogne T, Engstrøm AA, Jacobsen GW, Skranes J, Østgård HF, Martinussen M. Fetal growth, cognitive function, and brain volumes in childhood and adolescence. Obstet Gynecol. 2015;125:673-82.

In the univariable analyses, children born SGA had lower performance and full intelligence quotient (IQ) scores at five and nine years, and verbal IQ score at nine years compared with the control group. After adjusting for potential confounding factors, the IQ scores were comparable between the two groups, except for performance IQ score at five years (107.3 (95% CI 103.4, 111.1) vs. 112.5 (95% CI 109.9, 115.1)). While there were no differences in IQ scores between the SGA group without fetal growth restriction (FGR) and the control group, the SGA-FGR-group had lower performance IQ score at five years (103.5 (95% 95.9, 111.2) vs. 112.5 (95% CI 109.9, 115.1)) and nine years (96.2 (95% CI 86.0, 106.4) vs. 107.5 (95% CI 104.0, 110.9) compared with controls.

Children born SGA had smaller total intracranial volume at 15 years compared with controls, also after adjusting for confounding factors (1,472.5 cm^3 (95% CI 1,443.9, 1,501.0) vs. 1,548.4 cm^3 (95% CI 1,522.4, 1,574.4). There were no regional brain volumes differences between SGA children and controls after accounting for total intracranial volumes. However, there were differences between the SGA-FGR group and control group for thalamic volume (17.4 cm^3 (95% CI 16.5, 18.2) vs. 18.6 cm^3 (95% CI 18.3, 18.8)) and cerebellar white matter volumes (21.5 cm^3 (95%CI 20.1, 23.0) vs. 24.3 cm^3 (95% CI 23.8, 24.9)).

4.3 Paper III

Rogne T, Tielemans MJ, Chong MFF, Yajnik CS, Krishnaveni GV, Poston L, Jaddoe VWV, Steegers EAP, Joshi S, Chong YS, Godfrey KM, Yap FKP, Yahyaoui R, Thomas T, Hay G, Hogeveen M, Demir A, Saravanan P, Skovlund E, Martinussen MP, Jacobsen GW, Franco OH, Bracken MB, Risnes KR. Maternal vitamin B12 in pregnancy and risk of preterm birth and low birth weight: A systematic review and individual participant data meta-analysis. Am J Epidemiol (accepted for publication).

Six-hundred-six unique references were identified, of which 22 studies were eligible (11,993 pregnancies; Figure 4.1). Eighteen studies were included in the meta-analysis (11,216 pregnancies; 94% of all eligible pregnancies): Ten studies (8,928 pregnancies) provided individual patient data, two studies provided results from requested re-analyzes (973 pregnancies), and estimates from single studies were abstracted from the publications of six studies (1,315 observations).

Of the included studies, one was conducted in North America, nine in Europe, one in Africa, one in Oceania, and six in Asia. The number of pregnancies studied ranged from 84 to 5,641. Deficiency of vitamin B12 (B12) was identified in 0% to 69% of pregnancies (median 33%). The incidence of low birth weight (LBW) ranged from 0% to 33% (median 6%), preterm births from 4% to 14% (median 8%), and SGA from 5% to 32% (median 11%).

No linear association was observed between maternal vitamin B12 status and birth weight, but we observed an association between B12-deficiency and risk of giving birth to a LBW newborn (adjusted risk ratio (RR) 1.15 (95% CI 1.01, 1.31). For one standard deviation (SD) increase in B12 we observed an associated 11% reduced risk of preterm birth (95% CI 3, 18). Accordingly, B12-deficiency was associated with an increased risk of preterm birth (adjusted RR 1.21 (95% CI 0.99, 1.49). Maternal levels of B12 were not associated with birth weight SD scores (i.e. accounting for length of gestation and sex) or risk of SGA birth.

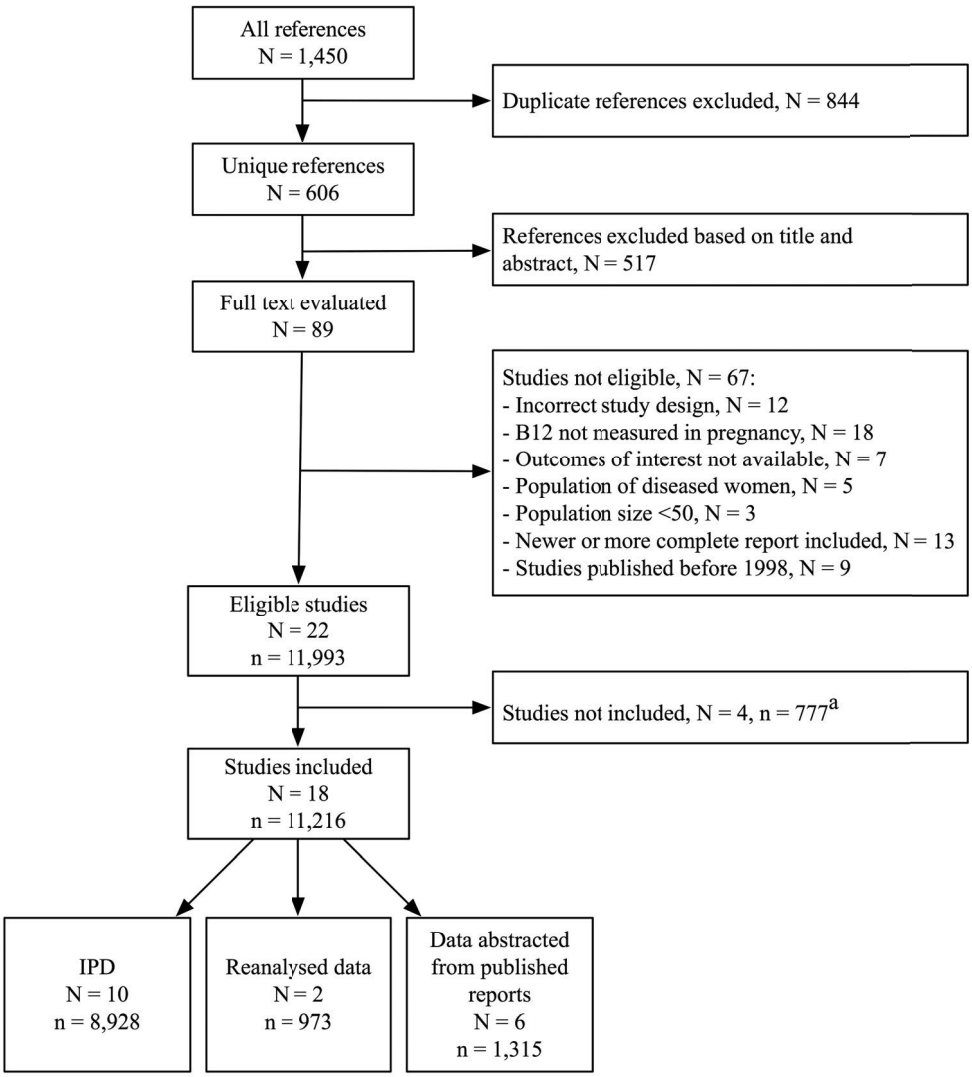


Figure 4.1 Flow chart of studies included in at least one of the meta-analyses of the association between vitamin B12 and birth weight or length of gestation. B12, vitamin B12; IPD, individual patient data; N, number of studies; n, number of pregnancies. a, IPD or re-analyses not provided, and results could not be abstracted from published reports

5. Discussion

5.1 Main findings

This thesis addressed how two nutritional factors during pregnancy may affect birth size: High glucose tolerance was found to be associated with impaired fetal growth, while low levels of vitamin B12 (B12) were associated with an increased risk of preterm birth, but not impaired fetal growth. Also, we observed reduced cognitive function and smaller regional brain volumes among children with small birth size due to impaired fetal growth, but not among the other children born small.

5.2 Epidemiological considerations

5.2.1 Causality

To identify *causes* of disease is ultimately the goal of much epidemiologic research. However, epidemiologists are usually reluctant to label their findings as causal, to much annoyance to the public. Scientific research is based on a refutationistic view, that is, we cannot prove a hypothesis, we can only discard it.[185] A much cited *viewpoint*[186] by Sir Austin Bradford Hill on causation has later been referred to as the *criteria for causation:*[187] strength, consistency, specificity, temporality, biological gradient, plausibility, coherence, experiment, and analogy. While details of the individual items will not be elaborated here, the general opinion is that a set of criteria for causality at best serves as guidance for critical appraisal.[187]

When determining causality, one will have to evaluate the *precision* of the results, as well as the *internal validity*. If one is to extrapolate findings in a paper to a population other than the study population, one also has to judge the *external validity* of the results. Each of these items will be addressed in the following sections.

5.2.2 Precision

In medical research we attempt to estimate the association between exposure and outcome. Due to random, unexplained variation between the exposed and unexposed, the estimated association always comes with some degree of uncertainty. If either exposure or outcome measure inaccurately, this will add to the random variation. With increasing number of subjects under study, one will expect this random variation to distribute equally among the exposed and unexposed. Precision, the converse of uncertainty, is often presented as a span of confidence around the mean (point estimate), usually as a 95% confidence interval (CI). By including more subjects under study, we will become more confident in our estimate, and the CI will narrow. There are nevertheless some important caveats to remember when appraising CIs.

First, the 95% CI reflects the range in which we are 95% confident that the *mean of the population* is found. That is, it does not predict the probability on an individual level. Furthermore, the CI only applies to populations equal to the one under study, as discussed in *5.2.4 External validity*. Lastly, the CI does not take bias into account (see next section). Thus, even with a narrow CI (i.e. precise estimate), we cannot conclude that an association is true. In technical terms, precision and CI deal with random error, but do not account for systematic error (i.e. bias).

There has been much debate about the use of "statistical significance" in statistics.[185] The term implies that statistical significant associations should be emphasized. The key statistic when assessing statistical significance is the *p*-value. When comparing an exposed and unexposed group, the interpretation of the *p*-value is as follows: If there is no difference between the groups being compared, what is the probability of finding the observed difference or a more extreme difference (even farther from the mean).[185] Statistical significance is assumed if this probability, the *p*-value, is below a certain cut-off, usually 5%. A *p*-value <5% equals a 95% CI that does not include the null-effect (i.e. no difference between the groups). With

increased power, the risk of chance findings is reduced. As discussed above, statistical significance does not take systematic error into account. Also, statistical significant results may be of little *clinical significance*, and vice versa. All papers in this thesis use 95% CIs to describe precision. "Statistical significance" and *p*-values have to some extent been used in *Paper I* and *Paper II.*

Paper I considered a fairly large sample of pregnancies, as compared with other studies with repeated ultrasound measurements during pregnancy. Due to recruitment of high risk pregnancies, we had a relatively large number of pregnancies with restricted fetal growth and small-for-gestational-age (SGA) births. Also, the oral glucose tolerance test (OGTT) is quite precise, as is estimation of fetal weight and newborn weight.[153,188] In terms of random error, we had a well powered study. As an example, we observed similar associations (i.e. effect estimates) between low blood glucose and SGA birth as another study, but our estimate was more precise.[189]

A different scenario was at play in *Paper II*. There were few observations in the main exposure groups; children born SGA with (n = 13) and without (n = 36) fetal growth restriction (FGR). Additionally, estimation of intelligence quotient (IQ) scores is subject to imprecision.[190] This left us with little statistical power. However, as the observed differences between the exposed and unexposed were considerable – that is, clinically significant – several of the associations were statistically significant.

Increased precision and statistical power is one of the alluring features of systematic reviews with meta-analyses. As elaborated in section *1.3.2 Systematic reviews*, one has to treat this increased statistical power with care. While randomized controlled trials ideally differ only by random error and are therefore suitable for inclusion in meta-analyses, observational studies are subject to bias. Because of this, some argue that one should generally refrain from doing meta-analyses in systematic reviews of observational studies.[34] In *Paper III* we made great

effort to collect raw data and ask for re-analyzed results in order to reduce bias (see *5.2.3 Internal validity*). The large number of observations allowed for precise estimates. For instance, a fairly large study (n = 986) reported an odds ratio for preterm birth of 0.81 (95% CI 0.64 to 1.03) per standard deviation (SD) increase of maternal B12, but concluded that there was "little or no" association between the two.[191] In a similar analysis, we observed a risk ratio (RR) for preterm birth of 0.89 (95% CI 0.82 to 0.97) per SD increase of maternal B12. Hence, even though the size of the effect estimate was somewhat smaller (but still clinically important), the increased precision enabled us to detect an association.

5.2.3 Internal validity

A precise estimate is of little value if it is false, i.e. biased. Although a myriad of specific biases have been identified, three main categories apply: selection bias, information bias, and confounding.[185] Each of these biases will be discussed in the setting of the thesis' three papers. Lastly, bias specific to systematic reviews will be discussed.

Selection bias

Selection bias arises when the subjects that are studied vary systematically from the subjects that are not studied.[185] The bias may occur both during the selection process and in the likelihood of being retained in the study.

Both the *Scandinavian SGA-study* and the studies included in the systematic review were prospective cohort studies where all study subjects that fulfilled pre-defined eligibility criteria were sought to be included in the studies. Whether the constructed eligibility criteria created a representative sample of the general population or not will be discussed further in section *5.2.4 External validity*. With respect to selection bias, we are concerned about failure to include eligible subjects, and loss to follow-up.

In the *Scandinavian SGA-study*, only 200 (3%) of the 5,922 eligible women recruited to the study failed to make first appointment. Half of these women failed to come to the first visit due to unknown reasons, the remainder due to social constraints of different sorts (e.g. "too time consuming").[55] There were 393 (20 %) drop outs, but they had the same risk status as the ones followed.[55]

In *Paper I*, the main reason for exclusion was that an OGTT had not been carried out. Among women that had done an OGTT, 70% were from the high-risk group (i.e. increased risk of giving birth to an SGA child), while there were 75% from the high-risk group among the women where an OGTT was not carried out. This imbalance may represent a slight selection bias. The OGTT was performed in gestational week 37. Women delivering preterm would for this reason not be included in *Paper I.* Risk factors for preterm birth, e.g. preeclampsia, may be subject to underrepresentation among those with an OGTT compared with those without an OGTT. One may speculate that if such unbalanced risk factors are associated with both the exposure (glucose tolerance) and outcome (fetal growth and birth weight) in *Paper I*, the results may be biased. The direction of the bias is uncertain.

There were 29 (13% of eligible) eligible children that were not followed in *Paper II*. Additionally, 84 (45% of followed) children were lost to follow-up from 5 to 15 years. Both may represent selection bias. Encouragingly, there were no notable differences in maternal or newborn characteristics between the three groups; eligible and followed, eligible but not followed, and eligible and lost to follow-up (Appendix).

All of the individual studies included in the systematic review recruited women irrespective of the risk of giving preterm birth or birth to a low birth weight (LBW) child. Still, low follow-up rate was a problem in many of the studies. The follow-up was adequate (data on exposure and outcome in >80% of eligible pregnancies) in only 8 of the 18 included studies (Appendix). Eight studies had a follow-up less than 80%, and the rate of follow-up was

uncertain in two additional studies. For each of the included studies, there may be varying reasons for loss to follow-up. If these are close to random, one would expect that the pooled effect estimate across all studies would be attenuated. If there are systematic reasons as to why women were lost to follow-up, bias may have been introduced. One such example could be that women with vitamin B12-deficiency developed anemia and withdrew from the study due to fatigue.

Information bias

If there is systematic error in the information collected about the study subjects, we have information bias.[185] We often structure exposures and outcomes into categories, and wrong classification due to information bias is often referred to as *misclassification bias.*[185] If the misclassification of the *outcome* is associated with the *exposure*, and vice versa, we have *differential misclassification.* Otherwise we refer to the misclassification as *non-differential.* Non-differential misclassification tend to attenuate the effect estimate. Differential misclassification, however, may bias the results in either direction, and is of particular concern.

We do not suspect that measurement of the outcomes depended on the exposure, glucose tolerance, in *Paper I*. Likewise, results from the glucose tolerance test were unlikely to depend on the ultrasound measurements of the fetus. Differential misclassification bias was therefore not suspected in *Paper I*. Non-differential misclassification bias is to some extent present in all epidemiological studies,[185] and it is likely that, for instance, some of the anthropometric measures of the fetus may have been erroneously registered.

Estimation of child IQ is not an exact science, and one may suspect that, in *Paper II*, the SGA-status may have affected how the IQ was scored. To limit such differential misclassification bias, the psychometrists assessing the children were blind to exposure status. In spite of that, one may not rule out that the psychometrists were influenced by certain traits

among children born SGA, e.g. short stature, perhaps by allowing lower scores among children suspected to be born SGA. In the unlikely event that SGA-status influenced scoring of IQ, the associations reported in *Paper II* may have been exaggerated. As with *Paper I*, non-differential misclassification may have occurred, particularly in less objective measures such as IQ scoring.

In *Paper III*, we found that all studies evaluated the outcomes of interest blind to exposure status in all but one study (Appendix). The single study in mention did not report how the outcomes were recorded. As vitamin B12 was ascertained prior to knowledge of the outcomes, we do not suspect differential misclassification bias to be present in the systematic review. Again, non-differential misclassification may have attenuated the results, e.g. by incorrect estimation of length of gestation.

Confounding

Confounding is a bewildering word for a simple matter; mixing of effects.[185] When you observe an association between an exposure and an outcome that is really due to an external factor (i.e. a "confounder"), we have *confounding*. A classic example is the apparent association between birth order and prevalence of Down's syndrome, which is confounded by maternal age.[192] The confounding factor must be *associated* with both the exposure and the outcome, but not be *caused* by either. Because the exposure groups compared in observational studies are not constructed at random (as in randomized controlled trials), confounding is close to inevitable.

While selection bias and information bias for the most part must be addressed when designing the study, confounding may also be accounted for when analyzing the data. This is done either by stratifying the dataset on the confounding variable(s), or by including the confounding variable(s) in the regression model. The latter technique has been applied in this thesis, with addition of some stratified analyses in *Paper III.* Confounding covariates should be identified by reasoning based on *a priori*, clinical knowledge, rather than associations identified

in statistical models.[185] In the absence of critical assessment of potential confounders, one will be at risk of either over-adjustment or missing key confounders.[170]

The *Scandinavian SGA-study* collected a large amount of data on each study participant. This enabled us to include the desired confounding variables in the regression models, as specified in *3.4.2 Multivariable analysis*. In *Paper I*, univariable and multivariable models yielded similar results. In *Paper II*, both children born SGA with FGR and those without FGR were associated with reduced IQ scores. However, after adjusting for maternal education and child sex, the notable associations were confined to children born SGA with FGR. Arguably some of the association between SGA and cognitive function in the univariable analyses was explained by maternal education or child sex. The results from the analyses of brain volumes were not affected by adjustment for confounding. Although we had many covariates to choose from in the *Scandinavian SGA-study*, it is difficult to take all potential confounders into account. Residual confounding may affect the associations reported in both *Paper I* and *Paper II*.

Control for confounding was more challenging in the systematic review. In traditional meta-analyses, the reviewer is not at liberty to define which confounders to include in the analyses. Indeed, most included studies will not have adjusted for the same confounding factors, if presenting adjusted models at all. By collecting raw data from the included studies, we had the opportunity to define which confounders to include in the models, and make sure the models were equal in all studies. Even with this flexibility, confounding may have affected our results. Only a few of the included studies provided data on all five identified confounding factors; maternal age, parity, body mass index (BMI) or weight, smoking and education. To balance adequate adjustment for confounding with a reasonable number of studies under consideration, we restricted our main multivariable model to adjustment for maternal age, parity and BMI or weight. Still, we found little discrepancy between the main model and the extended model in

the analyses of the association between B12 SD score and birth weight, and B12 SD score and risk of preterm birth (Appendix).

Maternal anthropometry was another example where we had to be pragmatic in *Paper III*. Both previous reports and observations from our systematic review have identified maternal overweight to be associated with lower blood levels of B12.[175,176] Failure to take maternal overweight into account would therefore underestimate an association between B12 and birth weight. While BMI is a more sensitive measure of adiposity than weight, most studies did not provide information on maternal BMI (or height and weight). To include this important confounding variable in our models, we adjusted for BMI when available, and allowed the remaining studies to adjusted for weight. This approach is feasible in meta-analyses using a two-step approach. Sensitivity analyses showed similar results in the analyses of birth weight and risk of preterm birth when adjusting for BMI in studies with data on both BMI and weight, as when adjusting for weight in the same studies (Appendix).

Bias in systematic reviews

In addition to bias that may affect within-study validity, bias may be present when synthesizing results across studies. A general discussion on how publication bias and reporting bias may distort the literature is found in the Introduction of this thesis (*1.3.2 Systematic reviews*).

We did not seek out unpublished studies for the systematic review. There is no general requirement to register longitudinal cohort studies ahead of conduct, as there is for clinical trials. We therefore deemed it unfeasible to identify all unpublished cohort studies that could have the necessary information on maternal vitamin B12 and birth weight or length of gestation. Of the 18 studies included, only ten studies reported to have objectives similar to those in our review. As follows, eight of the included studies had been published to report other associations than the ones under study in *Paper III*. Presumably, these eight studies may resemble the

unpublished literature on B12 and birth weight and length of gestation. In sensitivity analyses of B12 SD score in relation to birth weight and risk of preterm birth stratified by main objective (similar vs. different from our review's objectives), there were no differences in the results between the two strata in neither analysis (data not shown). Also, there was no suspicion of publication bias in the funnel plot of the association between B12 SD score in relation to birth weight. Based on these findings, we argue that there is little reason to suspect that publication bias has affected our results.

The reported results in the included studies did not allow for a traditional meta-analysis due to great heterogeneity in conduct and reporting of the analyses. Qualitative assessment of the association between B12 and birth weight was possible in 14 studies, but was inconclusive. Only two studies reported on the association between B12 and length of gestation, also inconclusive. However, by collecting individual patient data (IPD) and asking for results from re-analyzes, we were able to include 18 studies in the meta-analyses. Among the studies that provided either IPD or re-analyzed results (12 studies, 9,901 pregnancies), there was by design no reporting bias. Six studies (1,315 pregnancies) did not provide IPD or re-analyzed data, and were included in a restricted number of analyses, but their findings corroborated the findings of the main analyses. Four eligible studies (777 pregnancies) were not included in the review as they did not provide data, and had not reported the association of interest. Because these studies constituted a relatively small proportion of the eligible pregnancies (6%), it is unlikely that they biased our results. A telling example of how reporting bias may distort the literature is evident by comparing our review with a recently published systematic review with traditional meta-analysis (see *5.4.2 Systematic review*).[7]

Papers with significant findings have a greater tendency of being published in English, as compared to papers with negative results.[39] We applied no language restrictions in our

search. We also searched multiple databases applying a wide search string, thereby identifying studies irrespective of citation bias.

5.2.4 External validity

After considering the within-study validity, one must proceed further and consider to which populations the findings apply. In other words, the generalizability of the results. Findings valid only to the population under study is of little interest to the public. When designing a study, some argue that the study population should reflect the general population, so that the findings may be generalized to the population at large.[185] Others argue that a more selected study sample is preferable in order to evaluate causality.[185] The findings reported in the papers based on the *Scandinavian SGA-study* differ in their external validity as compared with the findings from the systematic review.

First considering *Paper I* and *Paper II*, there are several important differences between the population evaluated in the *Scandinavian SGA-study*, and the general population. The original study was designed in a time when genetics was associated with heredity, rather than DNA sequences: In order to evaluate "genetics", only parous women were recruited. To increase the number of events of the main outcome of interest, SGA birth, the longitudinal study was enriched with women at an increased risk of giving birth to an SGA child. Only pregnancies from Norway and Sweden were assessed. Furthermore, in the close to three decades that have passed since the enrolled women were followed during pregnancy, both the pregnant population and the antenatal care in Scandinavia have evolved.[193] Accordingly, multiple factors challenge the external validity of *Paper I* and *Paper II.*

In a sub-analysis restricted to the random sample (i.e. the population reference) in *Paper I*, we observed similar associations between glucose tolerance and fetal growth as we did in the main analysis. Encouragingly, associations akin to the ones observed in *Paper I* have been

reported in populations resembling the general pregnant population in both Japan[194] and Israel.[92] Still, this association was not observed in a general pregnant population in the USA.[189]

Again, in *Paper II*, the reported associations are found among offspring of previous high-risk pregnancies. For this paper, sub-analysis of only the random sample was not feasible due to few study subjects. As with the literature on glucose tolerance, both concordant and conflicting results have been found in study populations resembling the general populations in different countries.[139,140]

Paper I and *Paper II* describe associations that may or may not be applicable to the general population. Indeed, these are only single studies, and conflicting results have been published in multiple reports on multiple populations. The most sensible way to construe our findings from the *Scandinavian SGA-study* is to take note of the reported association, but keep in mind the population under study and the design of the studies. As have been discussed, individual studies should always be treated with caution (See *1.3.2 Systematic reviews*). In order for the topics covered in *Paper I* and *Paper II* to give guidance to clinical practice, one will have to conduct well designed systematic reviews, as discussed later (*5.5 Clinical implications and future studies*).

Paper III included close to all pregnancies evaluated in eligible reports on the association between vitamin B12 and length of gestation and birth weight. Furthermore, the included studies were conducted in 11 countries, in both high-income countries and low- and middle-income countries. Consequently, as compared to *Paper I* and *Paper II,* we have more information at hand in *Paper III* to evaluate the generalizability of the findings in this paper. Be that as it may, if the individual studies themselves did not evaluate representative samples of the pregnant population in the respective countries, we are no wiser. Based on the Newcastle-Ottawa scale, we found that 14 of the 18 included studies evaluated a "truly or somewhat representative sample of the pregnant population in the community" (Appendix).[149]

In the analysis of the association between maternal levels of B12 and *preterm birth*, there was little heterogeneity between the included studies, and the association was similar in high-income countries and low- and middle-income countries. Accordingly, this finding seems to be generalizable to pregnant populations. When stratifying the overall analysis of the association between maternal levels of B12 and *birth weight* into low- and middle-income countries and high-income countries, we observed an association in the former group of countries but not in the latter. Importantly, four of the five studies, and 95% of the pregnancies, included in the low- and middle-income group evaluated Indian populations. Indian pregnancies differ from pregnancies in other countries (high-income as well as low- and middle-income) due to a mainly vegetarian diet, making them susceptible to B12-deficiency.[195] Also, 30% of Indian newborns are born LBW, making them among the smallest in the world.[53] One may speculate that the discrepancy observed in the association between B12 and birth weight in high-income countries and low- and middle-income countries is due to different nutritional status and frequency of preterm births and LBW (see *5.3.2 Vitamin B12*).

Generalizability has in this section been treated on a population level, that is, whether the findings apply to a larger group of people. Yet, the findings may not apply to subgroups within these populations, even less so for individuals. It is important to remember that there is great disparity between findings on a population level, and the applicability on an individual level.

5.3 Appraisal of main findings

This thesis describes how two maternal factors – namely glucose tolerance and vitamin B12 levels in pregnancy – are associated with newborn size. Size at birth is a composite of fetal growth and length of gestation. Our results suggest that high glucose tolerance was associated with fetal growth, while B12 was more strongly associated with gestational length. These two

nutrients may affect fetal growth and length of gestation through separate processes: While glucose mediates its effect directly as an energy source and indirectly through stimulating production of growth factors such as insulin and insulin-like growth factors (IGF), vitamin B12 acts on the formation and methylation of DNA. The concept of "small at birth" was challenged in *Paper II*, where we found that fetal growth pattern in late pregnancy should be taken into account when assessing risk of impaired cognitive function.

5.3.1 Glucose tolerance

In *Paper I,* our approximation of glucose tolerance was the change in blood glucose from fasting state to two hours after an oral glucose challenge: delta (Δ) glucose. Low Δ glucose was chosen as the main exposure of interest partly based on previous reports on its association with restricted fetal growth.[196,197] The literature on high glucose tolerance or high insulin sensitivity in relation to fetal growth and birth weight is afflicted by varying definitions of the exposure. While insulin sensitivity is more closely related to the pathological mechanisms of gestational diabetes mellitus (GDM), glucose tolerance reflects the body's response to intake of glucose. Furthermore, the Δ glucose is more easily accessible than measures of insulin sensitivity. There is no commonly accepted definition of "too glucose tolerant"; indeed, there is no agreement whether such a condition is harmful during pregnancy. We decided to construct three groups based on Δ glucose; the upper and lower ten percent (*High* and *Low*, respectively), and the group of women between these extremes (*Medium*). This is an accepted way of categorizing groups.[198] Of the women in the *High*-group, 96% fulfilled the criteria for either GDM or overt diabetes mellitus (DM), which suggests that Δ glucose shows an adequate correlation with clinically familiar glucose metabolism disorders.

Several recent reports have evaluated the post-challenge glucose values from a glucose challenge test.[92,189,194] As two-hour glucose values show greater variation than do fasting

values, the two-hour values will contribute most to Δ glucose.[94] The group of pregnancies with low Δ glucose may therefore be compared with those with low post-challenge values in other reports.

In terms of glucose tolerance during pregnancy in relation to birth weight, the main focus has been on women with *intolerance*, that is, with GDM or DM. We found that women in the diabetogenic *High*-group carried heavier fetuses than in the two other groups, which is in accordance with the general understanding that GDM and DM are associated with accelerated fetal growth.[199]

Fewer studies have explored the association between proxies of increased glucose tolerance and fetal growth. A distinctive feature with our study is the longitudinal evaluation of fetal growth, where we observed a less favorable growth pattern in pregnancies with increased glucose tolerance. Most other studies, however, use measures at birth as substitute for fetal growth. Two early reports address the topic: The first study (1976) compared post-challenge hypoglycemic pregnancies with euglycemic pregnancies and observed a doubled risk for LBW births in hypoglycemic pregnancies.[200] A decade later (1986), a small, high-risk population, was evaluated and the researchers found a tenfold increased risk for SGA births in hypoglycemic pregnancies compared with controls.[201] The latter study compared hypoglycemic women with a combination of euglycemic and hyperglycemic women (i.e., not compared with women with "normal" glucose values), which may have exaggerated the effect estimate. More recent studies are discrepant. Some studies support the findings from the two old reports,[92,194,202,203] while others report no association between low post-challenge glucose values and restricted fetal growth.[189,204]

Women giving birth to growth restricted babies have repeatedly been found to have lower levels of glucose and insulin compared with controls.[205–207] Two small studies in the late 1990s observed an inverse association between insulin sensitivity and birth weight.[91,208] These

findings were later replicated in a much larger population.[93] We hypothesize than women in the *Low*-group were generally more insulin sensitive than women in the other two groups. If that theory is correct, our findings are in agreement with the three mentioned studies.

The underlying mechanism for an association between increased glucose tolerance and restricted fetal growth is uncertain. One explanation may be that low postprandial glucose values reflect pancreatic hyper-reactivity to glucose intake, reflected by over-secretion of insulin. Fasting and reduced dietary intake of glucose, on the other hand, has only a minor impact on circulating blood glucose concentrations.[5] Decreased levels of maternal glucose could lead to lower levels of glucose in the fetal bloodstream, and consequentially less stimulation of insulin and IGF production. GDM and DM is also associated with FGR (in addition to macrosomia), which may seem counterintuitive following the discussion above. Restricted fetal growth in pregnancies affected by GDM or DM may be caused be genetic mutations[120] or autoantibodies[87,209] that impair the fetal production or response to insulin. One may speculate that there is a common final pathway leading to FGR in pregnancies affected by high glucose tolerance and GDM or DM, namely impaired insulin production or sensitivity. Conversely, if low maternal glucose levels are not associated with decelerated fetal growth, this could be explained by a sufficient facilitated glucose transport across the placenta, even during periods of maternal hypoglycemia.

5.3.2 Vitamin B12

We found that low maternal levels of B12 in pregnancy was associated with an increased risk of preterm birth, especially among B12-deficient women. Among B12-deficient women we also observed an associated increased risk of LBW. The two contributing factors to a low birth weight are impaired fetal growth and being born preterm.[53] While there was an association between levels of B12 and risk of preterm birth, no such relation was found between B12 and

our proxies for fetal growth, i.e. birth weight SD score and SGA. This leads us to suspect that B12 may affect length of gestation, but not fetal growth rate.

Few studies have evaluated maternal B12 in relation to length of gestation. Indeed, only two of the identified eligible studies reported on this association, both observing no certain association.[191,210] In 2012, a case-control study reported *higher* levels of B12 among women delivering preterm than women delivering at term.[211] This apparent inverse association between B12 and length of gestation was recently reiterated in a semi-systematic review.[114] We know, however, that B12 declines during pregnancy.[112] Thus, one would expect that a longer length of gestation will coincide with a lower level of B12. For this reason, studies measuring B12 at birth were not included in our review. Consistent with our findings, a case-control study observed that low pre-pregnancy levels of B12 among Chinese women were associated with an increased risk of preterm birth, but not LBW or SGA.[212] The mechanism of how B12 may influence length of gestation is unclear. We know that low levels of vitamin B12 may cause an accumulation of homocysteine.[213] One pathway may be through hyperhomocysteinemia, which has been hypothesized to affect length of gestation through oxidative stress and placenta dysfunction.[214,215]

Preterm birth may be categorized into spontaneous and medically indicated, with varying etiologies.[70] Two important reasons for medically indicated preterm births are severe preeclampsia and suspected restricted fetal growth.[70,71] In case of the latter, we found no evidence of B12-deficiency to be associated with an increased risk of restricted fetal growth. The association between B12 and preeclampsia was not explored in our study, and remains unclear.[216–218] Increased levels of homocysteine, on the other hand, have been associated with an increased risk of preeclampsia.[216–219] Additionally, folate intake – which may lower circulating levels of homocysteine[220] – has been suggested to reduce risk of preeclampsia.[221] In terms of inflammation, one may hypothesize that low levels of B12, potentially through

increasing levels of homocysteine, may be associated with inflammation and premature rupture of membranes (PROM). Having said that, there does not seem to be a strong association between homocysteine and PROM.[222]

Although a sensitivity analysis stratified by type of preterm birth (spontaneous or medically indicated) would be informative, we did not have such information at hand. There are increasing numbers of medically indicated preterm births in high-income countries, while low- and middle-income countries generally have a lower rate of provider initiated preterm births.[73] We did not observe any clinical important difference in the association between B12-deficiency and risk of preterm birth among low- and middle-income countries as compared with high-income countries.

The mechanism behind a potential association between maternal B12 and fetal growth is uncertain. The suspected pathways have been those through DNA synthesis and methylation, the succinyl-coenzyme A pathway, and through accumulation of homocysteine.[114,213] High levels of homocysteine, hyperhomocysteinemia, has in epidemiological studies been identified as a risk factor for vascular disease.[223] Given the rich vasculature of the placenta, disorders related to placental dysfunction have also been investigated in the context of increased levels of homocysteine. A systematic review found hyperhomocysteinemia to be associated with increased risk of SGA births.[224] More recently, a Mendelian randomization study proposed that homocysteine was causally related to fetal growth.[195] Our review of the literature neither confirms nor refutes previous reports of an associations between B12 and fetal growth.

Because hyperhomocysteinemia in epidemiological studies has been associated with reduced birth weight, it is reasonable to evaluate the effect of lowering the levels of homocysteine. Yet, a Cochrane review of trials found no reduced risk of either LBW or preterm birth among women supplemented with folic acid compared with controls.[225] No RCTs of B12-supplementation during pregnancy were identified prior to the initiation of our review. Later,

two such trials have been published.[104,226] Intuitively, both studies observed an increase in levels of B12 in plasma among supplemented women compared with controls. In spite of that, there was no reduction in homocysteine levels. Risk of LBW tended to be lower in the supplemented group compared with the control group in both studies, but few events yielded imprecise results. Birth weight was comparable in the two groups in both studies. Additionally, B12-supplementation did not affect length of gestation or risk of preterm birth in either of the studies (C Duggan, Harvard University, personal communication, 2015).[104] Small but meaningful differences in risk of LBW and preterm birth may have been missed because the studies were too small; n = 256[226] and n = 68.[104]

Several non-systematic reviews on the association between B12 and birth weight have been published.[96,113,114] These narrative reviews suffer from bias detailed in *1.3.2 Systematic reviews*. Of particular notice, the three mentioned reviews assume an association between B12-deficiency and reduced fetal growth, which is evidence of file drawer bias, citation bias and selective reporting, rather than an explicitly evaluated association. The most recent narrative review also cited the previously mentioned case-control study that evaluated levels of B12 at birth among women delivering preterm compared with women delivering at term.[114] These reviews are at risk of spurious findings and are of limited value.

Of potentially greater value, a systematic review by Sukumar et al. published in April 2016 evaluated the association between maternal B12 and birth weight.[7] Nevertheless, that review had several important limitations that our review overcame. The greatest difference between that review and the review in this thesis was that we collected IPD for our review. A discussion of the findings of that review compared with those of our review is found under *Strengths and limitations* in section *5.4.2 Systematic review*. In short, they found no clear association between B12 and birth weight, and the association between B12 and length of gestation was not evaluated.

After conducting the final search in August 2016, at least two potentially eligible studies have been published. One of the studies, evaluating up to 496 pregnancies, reported no associations between maternal levels of B12 and length of gestation or birth weight.[227] A larger study (n = 4,114) observed a trend towards a positive association between B12 and length of gestation, but no association between B12 and birth weight.[127] Given the small size of the first study, and a trend in the second study similar to the one observed in our review, the conclusion of the systematic review would most likely not be altered by inclusion of these studies.

Paper I and *Paper III* were individual works, and we have made no attempts to evaluate glucose tolerance in relation to levels of vitamin B12. In addition to both being of importance to newborns there seems to be a close interplay between the two nutritional components. Experimental studies on sheep have found that limiting the fetal availability of B12 induces insulin resistance in the adult offspring.[228] Epidemiologic studies on humans have found vitamin B12-deficiency in pregnancy to be associated with insulin resistance in the offspring at the age of 6 years.[229] These associations have been explained by DNA methylation and epigenetic programming.[127,228,229] Furthermore, metformin, an anti-hyperglycemic therapy, have been found to reduce B12 concentrations.[230] The methionine-homocysteine-pathway, in which B12 is closely involved, is important for fetal myelination and neurodevelopment.[128,231]

5.3.3 Cognitive function and brain volumes

Cognitive function

Cognitive function of the children in the follow-up study of the *Scandinavian SGA-study* was explored in *Paper II*. SGA children had inferior results on crude performance IQ and full IQ at 5 and 9 years, and verbal IQ at 9 years compared with control children. After adjustment for maternal education, child sex and age at examination, there were no longer clear differences between the SGA group and the control group, except for performance IQ at 5 years. Several

factors may explain the change of results from univariable to multivariable analyses. Maternal education and child sex may independently explain some of the variation in IQ between the groups. Additionally, data on maternal education was missing among several subjects, resulting in a reduced number of observations under study, rendering type 2 error more likely.

A recent study similar to *Paper II* found a tendency of performance, verbal and full IQ scores among 7-year-old children born SGA to be about three points below that of their appropriate-for-gestational-age (AGA) peers.[232] Some earlier reports have found distinct differences between SGA children and controls,[138,139] while others have observed comparable scores.[141,233] Clinical and statistical heterogeneity between the studies may have contributed to the discrepant findings. In general, studies with the more conservative definitions of SGA tended to have the most pronounced differences. A systematic review published in 2016 observed an associated increased risk of lower cognitive scores in childhood and adolescence among children born SGA compared with controls.[234] Importantly, though, both publication bias and reporting bias are likely to have exaggerated the results.

We divided the group of children born SGA into those with and without FGR. The pathologically small children had substantially reduced performance IQ at 5 and 9 years compared with controls. This finding may suggest that impaired cognitive function among SGA children is greatest among those who also had signs of restricted growth in utero. Also, based on our results, children who were physiologically small at birth had only a slight – or no – reduction in cognitive abilities at 5 and 9 years compared with children of normal birth weight. A study from 2015 constructed exposure groups similar to those in *Paper II*.[235] They observed no reduced cognitive scores at 16 to 18 years among adolescents born SGA and FGR compared with controls. As we did, they defined FGR based on serial ultrasound measurements in late pregnancy. However, their cut-off for FGR was fetal growth below the 10th percentile, while

we used the 2.3rd percentile (i.e. 2 SD below the mean), which may partially explain some of the lack of association.

The verbal IQ did not differ substantially between the SGA-FGR group and the control group. This observation is in accord with the general understanding that verbal IQ to a great extent is influenced by environmental factors, while biological factors play a dominant role for the performance IQ.[236,237]

It is unknown which pathological processes among children born small may cause impaired cognitive function. As elaborated in *1.6 Consequences of small birth size*, it is unclear whether size at birth is the cause of later adverse outcomes, or if there are other factors associated with both birth size and the outcome. In *Paper II* we found that birth weight *per se* seemed to be of less importance than the fetal growth pattern in terms of later cognitive function. One may therefore speculate that causative factors are related to fetal health. There are several risk factors for FGR – maternal, placental and fetal – some of the more important ones being placental dysfunction, maternal smoking, and malnutrition.[238] Poor placenta function, for instance, may have deleterious effects on brain development.[239] Hence, we were interested to learn more about which brain structures were associated with FGR, and that might potentially explain the reduced cognitive scores among children born SGA and FGR.

Brain volumes

A global reduction in intracranial volume among children born small is caused by a combination of a compromised development of the gray and white matter. In late pregnancy, the gray and white matter are especially vulnerable.[125] Gray matter development is the leading determinant of total brain volume increase in late pregnancy, with a threefold absolute increase in volume from gestational week 30 to 40.[125] Glial proliferation and differentiation, and white matter formation starts in the second half of pregnancy.[124] Both *Paper II* and other studies have

found reduced total intracranial volume among children and adolescents born SGA.[132] In support of this association, induced restricted growth on fetal guinea-pigs cause reduced total brain weight.[240]

We observed symmetrical reduction in brain volumes among the physiologically small children, as has been described in previous studies of the same study population.[151,241] In the SGA-FGR group, however, normalized volumes of the thalamus and the cerebellar white matter were reduced compared with controls. We hypothesize that these structures are the most vulnerable to insults during pregnancy. In support of this hypothesis, a study on premature infants found the thalamus to be the supratentorial structure most commonly affected by neuronal loss and gliosis.[242]

Cell-to-cell interaction is imperative for neuronal and glial development.[243] Thalamus atrophy may be a result of either primary or secondary injury, or a combination of the two. The former would lead to secondary axonal injury, which in turn would result in hypomyelination and impaired development of other gray matter structures.[244] Secondary injury – that is, primary axonal injury – would potentially disrupt the connectivity between gray matter regions, and one would expect to find reduced volumes of the affected structures.[245,246] We can not infer which of the two pathways were at play in the SGA-FGR group, as the end result of primary and secondary thalamic injury is comparable in terms of volumes.

The reduced cerebellar white matter volume in the SGA-FGR group may be explained by its high metabolic activity in late pregnancy: First, the cerebellum has a faster growth rate than most brain structures in the second half of pregnancy.[247] Additionally, myelination of the cerebellar white matter starts in the third trimester.[248] Indeed, induced FGR in the second half of pregnancy resulted in reduced cerebellar white matter volumes in an experimental study on guinea-pigs.[144] As with the thalamus, reduced cerebellar volumes may be due to primary or secondary injury.

The reduced thalamic and cerebellar white matter volumes in the growth restricted SGA group may be independent events, or be correlated with one another. Studies on preterm infants have found cerebellar injuries to be associated with volume reductions in the cerebrum, and vice versa.[249,250] Interrupted cerebello-thalamo-cortical pathways has been hypothesized as the underlying mechanism. In support of this hypothesis, we also found reduced absolute volumes of the cerebral and cerebellar cortices in the SGA-FGR group. If the volume reductions were correlated, it is still open to question which of the structures that were the primary source of injury.

A limitation of the discussion above is that much of the current knowledge on perinatal neurodevelopment is based on research on offspring born preterm. Findings from these studies are not necessarily transferable to a term birth population, as some structural alterations are more common among preterm newborns (e.g. periventricular leukomalacia).[243]

Brain volumes and cognitive function

Both the thalamus and the cerebellum are considered important to cognitive function.[142,143] We were unable to identify specific nuclei within the thalamus. We hypothesize that the mediodorsal nucleus was affected as it is the most vulnerable nucleus in late pregnancy,[244] and it is associated with higher cognitive functions.[251] We furthermore speculate that impaired function of the cerebellar white matter, reflected by reduced volume, may interfere with the cerebellar cortex' contribution to cognitive function.

The lower IQ scores among the children in the SGA-FGR group may be explained by reduced thalamic and cerebellar white matter volumes. Yet, we found no correlation between any of the normalized brain structures and IQ. In general, there has been little success linking IQ scores to regional brain volumes, which may be explained by the complexity of human intelligence.[146] The volume *development* may prove more important to intellectual ability than

static measures, as has been reported to be the case for the cerebral cortex.[252,253] And there is more to brain function than regional volumes. A paper published last year found increased levels of glutamate and N-acetylaspartate to creatinine ratio in the frontal lobe among one-year-olds born SGA compared with controls.[254] The increased levels were associated with lower cognitive function and motor skills at two years of age.

5.4 Strengths and limitations

5.4.1 Scandinavian SGA-study

Two of the most valuable aspects of the *Scandinavian SGA-study* are the prospective design and comprehensive follow-up. Highly skilled midwives carried out the ultrasound measurements, and the fetal weights were estimated by a specifically adapted model. The repeated ultrasound measurements and extensive fetal weight data allowed us to identify FGR based on individual growth patterns, which is much preferred over cross-sectional measures.[255] Data on placenta function were not obtained, which could potentially have contributed with important information regarding the etiology of FGR.

The OGTT was performed on pregnant women regardless of risk factors for GDM, providing data on the whole spectrum of responses to a glucose challenge. We did, however, lack data on maternal insulin levels. This denied us from estimating insulin sensitivity, which would be a valuable exposure of interest in addition to Δ glucose. Also, we had no measure of fetal availability of glucose, which may depend on placenta function.[256]

In the follow-up study, all participants were assessed by a few, specially trained clinicians at the same study site through the entire duration of the study, minimizing inter-observer bias. Brain volumes of the adolescents were manually evaluated by a single expert in the field.

Paper I had a large study population. As detailed in *5.2.3 Internal validity*, loss to follow-up in *Paper II* was present, but unlikely to have introduced bias. Still, the reduced study population in *Paper II* due to attrition rendered a potentially underpowered study.

In *Paper II* we set the cut-off for FGR at 2 SD below the mean fetal growth of the control group. It may be argued that this cut-off was too conservative. As a result, borderline growth restricted children in the SGA non-FGR group may have contributed to a reduction of the overall IQ in that group. The choice was based on the assumption that the risk of adverse effects increases with an increasing negative deviation from expected fetal growth. Furthermore, minus 2 SD is the more common cut-off for suspected FGR,[257] and corresponds with a previous paper from the same study.[258]

One child with cerebral palsy (CP), probably due to severe birth asphyxia, was excluded from the SGA-FGR group in *Paper II*. The child had a serious brain injury and was a considerable negative outlier in terms of IQ scores and brain volumes. The group differences between the control and SGA-FGR groups were for this reason more pronounced with this subject included (Appendix). The etiology of CP has been debated.[259–261] Hypothetically, if the child's CP was solely a result of FGR, some of the effect of FGR on IQ and brain volumes would have been mediated through CP; excluding children with CP under such conditions would bias the results by allowing the SGA-FGR group to appear healthier than warranted.

5.4.2 Systematic review

A hallmark feature of the systematic review in this thesis is the use of IPD and re-analyzed data. Due to substantial heterogeneity in the published analyses, a traditional meta-analysis could not answer our research questions. The benefits of conducting IPD meta-analyses are readily illustrated by comparing our review with that of a recently published systematic review with traditional meta-analysis on the same topic.[7] As discussed in section *1.3.2 Systematic reviews*,

incomplete or selective reporting may reduce the replicability of studies and distort the literature.[22] The review by Sukumar et al. depended on reported associations, and were unable to include studies that had not reported their findings due to "insignificant results".[7] For instance, they excluded the study by Bergen et al., which contributed 5,641 observations in our review.[210] Of the studies included in the other review, we either included the same studies in our review, or excluded them based on pre-defined criteria. In one of the main analyses in the Sukumar review, they reported an adjusted OR of 1.70 (95% CI 1.16 to 2.50; I^2=84%) for LBW among B12-deficient women compared with non-deficient women. The direction of association was similar to a comparable analysis in our review, RR 1.15 (95% CI 1.01, 1.31; I^2=5%), but of greater magnitude. The result was probably skewed due to reporting bias. Additionally, of eight individual results included in Sukumar's meta-analysis, five evaluated most of the same women from a single original study, exaggerating the influence of a single, outlying study.[100,262] Our review permitted an unbiased summary of the published literature as we included 94% of all eligible participants. We were able to include ten times as many pregnancies in our analysis compared with Sukumar's analysis.

Another strength of our review, allowed for by collection of IPD and re-analyzed data, was the low heterogeneity. The I^2 in our primary analyses ranged from 0% to 30%. As comparison, the I^2 in the main analyses in the review by Sukumar et al. ranged from 74% to 98%. It is widely accepted that given high heterogeneity, one should refrain from doing a meta-analysis.[263] In the Sukumar analysis of low B12 and LBW, I^2 was 84%: The high heterogeneity arose from varying comparison groups (e.g. maternal blood and cord blood), outcome groups (e.g. LBW and SGA), and confounders included in the multivariable models. We were able to conduct the analyses exactly the same way in all studies, adjusting for the same important confounder, providing an I^2 of 5% in the equivalent analysis.

While Sukumar et al. had to construct separate analyses based on the availability of data, our review included all studies with IPD or results from re-analyses in all the main analyses. Additionally, we were able to perform a great number of sensitivity analyses, both on study level (e.g. country) and on individual level (e.g. overweight) factors. Sukumar et al. did not have this opportunity.

Another strength was that our analyses were not post-hoc, but followed a detailed protocol. Our rigorous inclusion criteria allowed for valid analyses. For instance, we did not include studies that measured B12 at delivery, as this would have biased the results towards an inverse association between B12 and length of gestation (*5.3.2 Vitamin B12*).

There are several limitations. We decided not to consider unpublished studies for our review. Also, we were unable to include four eligible studies in our analyses. However, it is unlikely that these limitations have biased our results (see *5.2.3 Internal validity*).

Due to our selected number of outcomes of interest, we were unable to explore underlying factors for preterm birth, as described in *5.3.2 Vitamin B12*.

Finally, given that the review was based on observational studies, confounding factors were most likely at play. Although we found little discrepancy in the pooled results of adjusted main models as compared to extended adjusted models (*5.2.3 Internal validity*), the reported association between B12-deficiency and risk of preterm birth should be interpreted with caution. Notably, B12-deficiency may be a proxy for inadequate nutritional status, and it is possible that some of our findings are related to nutritional status, not specifically to B12. A vegan or predominantly plant-based diet is low in B12. Such a diet is frequently also low in other nutrients, such as iodine, vitamin D, zinc, iron, riboflavin, as well as protein and energy.[264] In our review, we did not have the necessary data to stratify our analyses by dietary intake of these nutrients. Some of these nutrients, such as vitamin D and zinc, have been associated with risk of preterm birth.[265,266] It is therefore possible that the finding of increased risk of preterm

birth with low B12 status is actually a result of mixed nutritional inadequacy or even deficiency in another nutrient than B12 or energy.

5.5 Clinical implications and future studies

It is tempting to rush from reporting associations in publications to recommending changes in clinical practice. It is nevertheless important to restrain oneself from making unwarranted extrapolations. Remember, most research findings reach false conclusions (*1.3 Evidence-based medicine*).[13] For instance, a systematic review with meta-analysis found a "causal" relation between increased homocysteine levels and risk of myocardial infarction.[267] Instead of recommending supplementation of homocysteine-lowering micronutrients, the researchers duly recommended conduction of trials. A large, simple RCT was conducted and, to much surprise, supplementation of homocysteine-lowering B-vitamins tended to *increase* the risk of myocardial infarction.[268] So, while changes in clinical practice will have to wait, what follows is a discussion of the next steps for future studies.

The level of evidence of the papers in this thesis varies. The two first publications are single, observational studies, while the last publication is a synthesis of close to all publications on a single topic. This affects what types of studies remain to be done.

We found a tendency of reduced fetal growth among the most glucose tolerant pregnant women in *Paper I*. To the best of my knowledge, no systematic reviews have been conducted exploring the association between high glucose tolerance during pregnancy and birth weight. This will be the next logical step to reach more valid evidence. As there has been much focus on DM, GDM and glucose tolerance during pregnancy, one may expect that a large number of studies have data on glucose tolerance and birth weight. Thus, there is opportunity to be more specific of which studies to include; e.g. only include studies that have measured both glucose

and insulin prior to and after a glucose challenge. Collection of IPD and re-analyzed data will probably be necessary.

The novel findings in *Paper II* were that impaired cognitive function and reduced regional brain volumes were confined to the children born small due to restricted fetal growth. As opposed to the topic of *Paper I*, there have been published several systematic reviews relating size at birth with later cognitive function.[234,269,270] Still, these reviews did not account for the longitudinal growth pattern. In order to appropriately do so, it will most likely be necessary to collect IPD from studies that have the necessary information; repeated ultrasound measurements during pregnancy, and cognitive tests and/or regional brain volume measurements during childhood and adolescence. If robust evidence is provided that mainly restricted fetal growth, and not small size at birth, is important for later cognitive function, this may have important implications. For instance, early intervention programs on newborns born preterm seem to improve cognitive function in pre-school children.[271] This kind of intervention may be useful for newborns born small due to FGR.

Paper III is closer to have a clinical impact than the two first publications of this thesis. One would expect from our findings that increasing the levels of B12 in a pregnant woman would reduce her risk of delivering preterm. However, as have been elaborated previously in this thesis, the reported associations do not prove a causal link between B12 levels in the mother and preterm birth. The World Health Organization (WHO) does not recommend B12-supplementation in pregnancy in low- and middle-income countries.[103] As of today, we know very little of the effect of B12-supplementation in pregnancy, but it is practiced even so. Two small trials have already been conducted on the topic.[104,226] A recent Cochrane review encourage that the effect of single micronutrient supplementation (in addition to iron and folic acid) in pregnancy should be explored.[78] *Paper III* provides robust evidence that support trials on B12-supplementation in pregnancy. Considering the pattern of evidence to be provided to

change clinical practice guidelines, a systematic review of trials on B12-supplementation should be conducted after initial trials. Depending on the findings of that systematic review, B12-supplementation should either be encouraged or discouraged during pregnancy. The dosage of B12 in the multiple vitamin and mineral supplement by WHO, United Nations Children's Fund and the World Food Programme, may also be increased depending on future results.[103]

The above-mentioned recommendations for future studies include systematic reviews as an integral step towards evidence-based medical practice. I am of the strong opinion that high quality systematic reviews are essential to reduce research waste and improve clinical practice. Conduct of such studies should be encouraged, not only to research fellows within the field of public health, but to researchers in all fields.

6. Conclusion

The findings in this thesis show that both high glucose tolerance and low maternal levels of vitamin B12 during pregnancy may affect birth size: The former by impaired fetal growth, and the latter by shorter length of gestation. Additionally, the interpretation of "small size at birth" has been challenged: In terms of brain development and cognitive function in childhood and adolescence, adverse outcomes seem to be greatest among those born small at term due to impaired fetal growth. The literature on high glucose tolerance in pregnancy in relation to birth weight, and on fetal growth restriction in relation to cognitive function and regional brain volumes needs to be systematically reviewed. Randomized controlled trials of vitamin B12-supplementation during pregnancy are also strongly encouraged.

References

1 Barker DJ. The fetal and infant origins of adult disease. *BMJ* 1990; **301**: 1111.

2 Kramer MS. Determinants of low birth weight: methodological assessment and meta-analysis. *Bull World Health Organ* 1987; **65**: 663–737.

3 Risnes KR, Vatten LJ, Baker JL, *et al.* Birthweight and mortality in adulthood: a systematic review and meta-analysis. *Int J Epidemiol* 2011; **40**: 647–61.

4 Ramakrishnan U, Grant F, Goldenberg T, Zongrone A, Martorell R. Effect of women's nutrition before and during early pregnancy on maternal and infant outcomes: A systematic review. *Paediatr Perinat Epidemiol* 2012; **26**: 285–301.

5 Harding JE. The nutritional basis of the fetal origins of adult disease. *Int J Epidemiol* 2001; **30**: 15–23.

6 Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med* 1998; **15**: 539–53.

7 Sukumar N, Rafnsson S, Kandala N, Bhopal R, Yajnik C, Saravanan P. Prevalence of vitamin B-12 insufficiency during pregnancy and its effect on offspring birth weight: a systematic review and meta-analysis. *Am J Clin Nutr* 2016; **103**: 1232–51.

8 Wilcox AJ. On the importance--and the unimportance--of birthweight. *Int J Epidemiol* 2001; **30**: 1233–41.

9 Chalmers I, Bracken MB, Djulbegovic B, *et al.* How to increase value and reduce waste when research priorities are set. *Lancet* 2014; **383**: 156–65.

10 Lawn JE, Blencowe H, Oza S, *et al.* Every Newborn: progress, priorities, and potential beyond survival. *Lancet* 2014; **384**: 189–205.

11 Norwegian Institute of Public Health. [Register of births in Norway in 2014]. 2016.

12 UNICEF, WHO, World Bank, UN-DESA Population Division. Levels and trends in child mortality 2015. 2015.

13 Ioannidis JP a. Why most published research findings are false. *PLoS Med* 2005; **2**: e124.

14 Straus SE, Glasziou P, Richardson WS, Haynes RB. Evidence-based medicine: How to practice and teach it, Fourth Edi. London: Churchill-Livingstone, 2011.

15 Sackett DL. Evidence-based medicine. *Semin Perinatol* 1997; **21**: 3–5.

16 Glasziou P, Burls A, Gilbert R. Evidence based medicine and the medical curriculum. *BMJ* 2008; **337**: a1253.

17 Ulvenes LV, Aasland O, Nylenna M, Kristiansen IS. Norwegian physicians' knowledge of and opinions about evidence-based medicine: Cross-sectional study. *PLoS One* 2009; **4**: 12–5.

18 Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet* 2009; **374**: 86–9.

19 Ioannidis JPA, Greenland S, Hlatky MA, *et al.* Increasing value and reducing waste in research design, conduct, and analysis. *Lancet* 2014; **383**: 166–75.

20 Al-Shahi Salman R, Beller E, Kagan J, *et al.* Increasing value and reducing waste in biomedical research regulation and management. *Lancet* 2014; **383**: 176–85.

21 Chan AW, Song F, Vickers A, *et al.* Increasing value and reducing waste: Addressing inaccessible research. *Lancet* 2014; **383**: 257–66.

22 Glasziou P, Altman DG, Bossuyt P, *et al.* Reducing waste from incomplete or unusable reports of biomedical research. *Lancet* 2014; **383**: 267–76.

23 Bastian H, Glasziou P, Chalmers I. Seventy-five trials and eleven systematic reviews a day: How will we ever keep up? *PLoS Med* 2010; **7**. DOI:10.1371/journal.pmed.1000326.

24 Jacobsen GW, Nylenna M. [Co-publication of systematic reviews]. *Tidsskr Den Nor legeforening* 2016; **136**: 376.

25 Bero LA, Grilli R, Grimshaw JM, *et al.* Getting research findings into practice: Closing the gap between research and practice: an overview of systematic reviews of interventions to promote the implementation of research findings. *Bmj* 1998; **317**: 465–8.
26 The Cochrane Collaboration. Cochrane Library, Glossary. 2016. http://community-archive.cochrane.org/glossary (accessed Feb 1, 2016).
27 Pearson K. Report on certain enteric fever inoculation statistics. *BMJ* 1904; **3**: 1243–6.
28 Island R. Climbing a Medical Everest. *Science (80-)* 2003; **300**: 2024–5.
29 Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med* 1992; **327**: 248–54.
30 Moynihan R, Bero L, Ross-Degnan D, *et al.* Coverage by the news media of the benefits and risks of medications. *N Engl J Med* 2000; **342**: 1645–50.
31 Jannot AS, Agoritsas T, Gayet-Ageron A, Perneger T V. Citation bias favoring statistically significant studies was present in medical research. *J Clin Epidemiol* 2013; **66**: 296–301.
32 Dickersin K. How important is publication bias? A synthesis of available data. *AIDS Educ Prev* 1997; **9**: 15–21.
33 Rosenthal R. The file drawer problem and tolerance for null results. *Psychol Bull* 1979; **86**: 638–41.
34 Egger M, Schneider M, Davey Smith G. Spurious precision? Meta-analysis of observational studies. *BMJ* 1998; **316**: 140–4.
35 Chavalarias D, Ioannidis JPA. Science mapping analysis characterizes 235 biases in biomedical research. *J Clin Epidemiol* 2010; **63**: 1205–15.
36 Brind J, Chinchilli VM, Severs WB, Summy-Long J. Induced abortion as an independent risk factor for breast cancer: a comprehensive review and meta-analysis. *J Epidemiol Community Health* 1996; **50**: 481–96.
37 Melbye M, Wohlfahrt J, Olsen JH, *et al.* Induced abortion and the risk of breast cancer. *N Engl J Med* 1997; **336**: 81–5.
38 Kirkham JJ, Dwan KM, Altman DG, *et al.* The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. *BMJ* 2010; **340**: c365.
39 Egger M, Smith GD, Altman DG. Systematic Reviews in Health Care, Second. BMJ Publishing Group, 2001.
40 Dwan K, Altman DG, Arnaiz J a, *et al.* Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS One* 2008; **3**: e3081.
41 Dal-Ré R, Ioannidis JP, Bracken MB, *et al.* Making prospective registration of observational research a reality. *Sci Transl Med* 2014; **6**: 224cm1.
42 Moher D, Glasziou P, Chalmers I, *et al.* Increasing value and reducing waste in biomedical research: Who's listening? *Lancet* 2015; **6736**: 1–15.
43 Bracken M. Preregistration of Epidemiology Protocols. *Epidemiology* 2011; **22**: 607.
44 Liberati A, Altman DG, Tetzlaff J, *et al.* The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009; **339**: b2700.
45 Vandenbroucke JP, von Elm E, Altman DG, *et al.* Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS Med* 2007; **4**: e297.
46 Stewart L., Parmar MK. Meta-analysis of the literature or of individual patient data: is there a difference? *Lancet* 1993; **341**: 418–22.
47 Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: rationale, conduct, and reporting. *BMJ* 2010; **340**: c221.

48 Vale CL, Rydzewska LHM, Rovers MM, Emberson JR, Gueyffier F, Stewart L a. Uptake of systematic reviews and meta-analyses based on individual participant data in clinical practice guidelines: descriptive study. *Bmj* 2015; **350**: h1088–h1088.
49 Smith GCS, Pell JP. Parachute use to prevent death and major trauma related to gravitational challenge: systematic review of randomised controlled trials. *BMJ* 2003; **327**: 1459–61.
50 Mickan S, Burls A, Glasziou P. Patterns of 'leakage' in the utilisation of clinical guidelines: a systematic review. *Postgrad Med J* 2011; **87**: 670–9.
51 Susser M, Marolla FA, Fleiss J. Birth weight, fetal age and perinatal mortality. *Am J Epidemiol* 1972; **96**: 197–204.
52 Wilcox a J, Skjaerven R. Birth weight and perinatal mortality: the effect of gestational age. *Am J Public Health* 1992; **82**: 378–82.
53 United Nations Children's Fund and World Health Organization. Low Birthweight: Country, regional and global estimates. 2004.
54 Copper RL, Goldenberg RL, Creasy RK, *et al.* A multicenter study of preterm birth weight and gestational age-specific neonatal mortality. *Am J Obstet Gynecol* 1993; **168**: 78–84.
55 Bakketeig LS, Jacobsen G, Hoffman HJ, *et al.* Pre-pregnancy risk factors of small-for-gestational age births among parous women in Scandinavia. *Acta Obstet Gynecol Scand* 1993; **72**: 273–9.
56 Urquia ML, Ray JG. Seven caveats on the use of low birthweight and related indicators in health research. *J Epidemiol Community Health* 2012; **66**: 971–5.
57 Ray JG, Sgro M, Mamdani MM, *et al.* Birth weight curves tailored to maternal world region. *J Obstet Gynaecol Can* 2012; **34**: 159–71.
58 Wilcox AJ. Intrauterine growth retardation: beyond birthweight criteria. *Early Hum Dev* 1983; **8**: 189–93.
59 Anderson NH, Sadler LC, Mckinlay CJ, Mccowan LME. INTERGROWTH 21st versus customized birthweight standards for identification of perinatal mortality and morbidity. *Am J Obstet Gynecol* 2015; **214**: 509.e1–509.e7.
60 Carberry AE, Gordon A, Bond DM, Hyett J, Raynes-Greenow CH, Jeffery HE. Customised versus population-based growth charts as a screening tool for detecting small for gestational age infants in low-risk pregnant women. Cochrane database Syst. Rev. 2014; **5**: CD008549.
61 Pardi G, Marconi AM, Cetin I. Placental-fetal interrelationship in IUGR fetuses - A review. *Placenta* 2002; **23 Suppl A**: S136–41.
62 Berkowitz GS, Chitkara U, Rosenberg J, *et al.* Sonographic estimation of fetal weight and Doppler analysis of umbilical artery velocimetry in the prediction of intrauterine growth retardation: a prospective study. *Am J Obstet Gynecol* 1988; **158**: 1149–53.
63 Cnossen JS, Morris RK, ter Riet G, *et al.* Use of uterine artery Doppler ultrasonography to predict pre-eclampsia and intrauterine growth restriction: a systematic review and bivariable meta-analysis. *CMAJ* 2008; **178**: 701–11.
64 Olsen J. David Barker (1938-2013) - A giant in reproductive epidemiology. *Acta Obstet Gynecol Scand* 2014; **93**: 1077–80.
65 Nyberg D a, Abuhamad A, Ville Y. Ultrasound assessment of abnormal fetal growth. *Semin Perinatol* 2004; **28**: 3–22.
66 Bamberg C, Kalache KD. Prenatal diagnosis of fetal growth restriction. *Semin Fetal Neonatal Med* 2004; **9**: 387–94.
67 Haram K, Søfteland E, Bukowski R. Intrauterine growth restriction: effects of physiological fetal growth determinants on diagnosis. *Obstet Gynecol Int* 2013; **2013**: 1–9.

68 Lin CC, Santolaya-Forgas J. Current concepts of fetal growth restriction: part I. Causes, classification, and pathophysiology. *Obstet Gynecol* 1998; **92**: 1044–55.
69 Brodsky D, Christou H. Current concepts in intrauterine growth restriction. *J Intensive Care Med* 2004; **19**: 307–19.
70 Goldenberg RL, Culhane JF, Iams JD, Romero R. Epidemiology and causes of preterm birth. *Lancet* 2008; **371**: 75–84.
71 Gyamfi-Bannerman C, Fuchs KM, Young OM, Hoffman MK. Nonspontaneous late preterm birth: Etiology and outcomes. *Am J Obstet Gynecol* 2011; **205**: 456.e1–456.e6.
72 Gardosi JO. Prematurity and fetal growth restriction. In: Early Human Development. 2005: 43–9.
73 Blencowe H, Cousens S, Oestergaard MZ, *et al.* National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *Lancet* 2012; **379**: 2162–72.
74 Brooks a. a., Johnson MR, Steer PJ, Pawson ME, Abdalla HI. Birth weight: Nature or nurture? *Early Hum Dev* 1995; **42**: 29–35.
75 Lunde A, Melve KK, Gjessing HK, Skjaerven R, Irgens LM. Genetic and Environmental Influences on Birth Weight, Birth Length, Head Circumference, and Gestational Age by Use of Population-based Parent-Offspring Data. *Am J Epidemiol* 2007; **165**: 734–41.
76 Lumey LH, Ravelli AC, Wiessing LG, Koppe JG, Treffers PE, Stein ZA. The Dutch famine birth cohort study: design, validation of exposure, and selected characteristics of subjects after 43 years follow-up. *Paediatr Perinat Epidemiol* 1993; **7**: 354–67.
77 Barker DJP. Intrauterine programming of adult disease. *Mol Med Today* 1995; **1**: 418–23.
78 Haider BA, Bhutta ZA. Multiple-micronutrient supplementation for women during pregnancy. *Cochrane database Syst Rev* 2015; **11**: CD004905.
79 Freinkel N. Banting Lecture 1980. Of pregnancy and progeny. *Diabetes* 1980; **29**: 1023–35.
80 Scholl TO, Sowers M, Chen X, Lenders C. Maternal glucose concentration influences fetal growth, gestation, and pregnancy complications. *Am J Epidemiol* 2001; **154**: 514–20.
81 Persson B, Hansson U. Hypoglycaemia in pregnancy. *Baillieres Clin Endocrinol Metab* 1993; **7**: 731–9.
82 Lockitch G. Clinical biochemistry of pregnancy. *Crit Rev Clin Lab Sci* 1997; **34**: 67–139.
83 Oliver MH, Harding JE, Breier BH, Evans PC, Gluckman PD. Glucose but not a mixed amino acid infusion regulates plasma insulin-like growth factor-I concentrations in fetal sheep. *Pediatr Res* 1993; **34**: 62–5.
84 Laviola L, Perrini S, Belsanti G, *et al.* Intrauterine growth restriction in humans is associated with abnormalities in placental insulin-like growth factor signaling. *Endocrinology* 2005; **146**: 1498–505.
85 Wilkin TJ, Murphy MJ. The gender insulin hypothesis: why girls are born lighter than boys, and the implications for insulin resistance. *Int J Obes (Lond)* 2006; **30**: 1056–61.
86 Metzger BE. Hyperglycemia and Adverse Pregnancy Outcome (HAPO) Study: associations with neonatal anthropometrics. *Diabetes* 2009; **58**: 453–9.
87 Li Z, Wu TM, Ming WJ, Chen X, Xiao XM. The influence of maternal islet beta-cell autoantibodies in conjunction with gestational hyperglycemia on neonatal outcomes. *PLoS One* 2015; **10**: 1–8.
88 Lowe LP, Coustan DR, Metzger BE, *et al.* Hyperglycemia and Adverse Pregnancy Outcome (HAPO) Study: Associations of maternal A1C and glucose with pregnancy outcomes. *Diabetes Care* 2012; **35**: 574–80.

89 Yang X, Hsu-Hage B, Zhang H, Zhang C, Zhang Y, Zhang C. Women with impaired glucose tolerance during pregnancy have significantly poor pregnancy outcomes. *Diabetes Care* 2002; **25**: 1619–24.
90 Steegers EAP, Von Dadelszen P, Duvekot JJ, Pijnenborg R. Pre-eclampsia. *Lancet* 2010; **376**: 631–44.
91 Catalano PM, Drago NM, Amini SB. Maternal carbohydrate metabolism and its relationship to fetal growth and body composition. *Am J Obs Gynecol* 1995; **172**: 1464–70.
92 Melamed N, Hiersch L, Peled Y, Hod M, Wiznitzer A, Yogev Y. The association between low 50 g glucose challenge test result and fetal growth restriction. *J Matern Fetal Neonatal Med* 2013; **26**: 1107–11.
93 Dalfrà MG, Pacini G, Parretti E, Ragazzi E, Mello G, Lapolla A. Elevated insulin sensitivity and β-cell function during pregnancy in mothers of growth-restricted newborns. *Am J Physiol Endocrinol Metab* 2011; **301**: E25–30.
94 Bartoli E, Fra GP, Carnevale Schianca GP. The oral glucose tolerance test (OGTT) revisited. *Eur J Intern Med* 2011; **22**: 8–12.
95 Muniyappa R, Lee S, Chen H, Quon MJ. Current approaches for assessing insulin sensitivity and resistance in vivo: advantages, limitations, and appropriate usage. *Am J Physiol Endocrinol Metab* 2008; **294**: E15–26.
96 Rush EC, Katre P, Yajnik CS. Vitamin B12: one carbon metabolism, fetal growth and programming for chronic disease. *Eur J Clin Nutr* 2014; **68**: 2–7.
97 Allen RH, Stabler SP, Savage DG, Lindenbaum J. Metabolic abnormalities in cobalamin (vitamin B12) and folate deficiency. *FASEB J* 1993; **7**: 1344–53.
98 Chiang PK, Gordon RK, Tal J, *et al.* S-Adenosylmethionine and methylation. *FASEB J* 1996; **10**: 471–80.
99 Halarnkar PP, Blomquist GJ. Comparative aspects of propionate metabolism. *Comp Biochem Physiol B* 1989; **92**: 227–31.
100 Muthayya S, Kurpad A, Duggan C, *et al.* Low maternal vitamin B12 status is associated with intrauterine growth retardation in urban South Indians. *Eur J Clin Nutr* 2006; **60**: 791–801.
101 Ray JG, Goodman J, O'Mahoney PR a, Mamdani MM, Jiang D. High rate of maternal vitamin B12 deficiency nearly a decade after Canadian folic acid flour fortification. *QJM* 2008; **101**: 475–7.
102 Pathak P, Kapil U, Yajnik CS, Kapoor SK, Dwivedi SN, Singh R. Iron, folate, and vitamin B12 stores among pregnant women in a rural area of Haryana State, India. *Food Nutr Bull* 2007; **28**: 435–8.
103 Gernand AD, Schulze KJ, Stewart CP, Jr KPW. Micronutrient deficiencies in pregnancy worldwide: health effects and prevention. *Nat Rev Endocrinol* 2016; **12**: 274–89.
104 Siddiqua TJ, Ahmad SM, Ahsan KB, *et al.* Vitamin B12 supplementation during pregnancy and postpartum improves B12 status of both mothers and infants but vaccine response in mothers only: a randomized clinical trial in Bangladesh. *Eur J Nutr* 2016; **55**: 281–93.
105 Allen LH. Vitamin B12 metabolism and status during pregnancy, lactation and infancy. *Adv Exp Med Biol* 1994; **352**: 173–86.
106 Yajnik CS, Deshpande SS, Lubree HG, *et al.* Vitamin B12 deficiency and hyperhomocysteinemia in rural and urban Indians. *J Assoc Physicians India* 2006; **54**: 775–82.
107 Yakut M, Ustün Y, Kabaçam G, Soykan I. Serum vitamin B12 and folate status in patients with inflammatory bowel diseases. *Eur J Intern Med* 2010; **21**: 320–3.
108 Stabler SP, Allen RH. Vitamin B12 deficiency as a worldwide problem. *Annu Rev Nutr*

2004; **24**: 299–326.
109 Obeid R, Morkbak AL, Munz W, Nexo E, Herrmann W. The cobalamin-binding proteins transcobalamin and haptocorrin in maternal and cord blood sera at birth. *Clin Chem* 2006; **52**: 263–9.
110 Dror DK, Allen LH. Interventions with vitamins B6, B12 and C in pregnancy. *Paediatr Perinat Epidemiol* 2012; **26 Suppl 1**: 55–74.
111 Bae S, West AA, Yan J, *et al.* Vitamin B-12 status differs among pregnant, lactating, and control women with equivalent nutrient intakes. *J Nutr* 2015; **145**: 1507–14.
112 Milman N, Bergholt T, Byg K-E, Eriksen L, Hvas A-M. Reference intervals for haematological variables during normal pregnancy and postpartum in 434 healthy Danish women. *Eur J Haematol* 2007; **79**: 39–46.
113 Pepper MR, Black MM. B12 in fetal development. *Semin Cell Dev Biol* 2011; **22**: 619–23.
114 Finkelstein JL, Layden AJ, Stover PJ. Vitamin B-12 and Perinatal Health. *Adv Nutr* 2015; **6**: 552–63.
115 Hogeveen M, Blom HJ, van der Heijden EH, *et al.* Maternal homocysteine and related B vitamins as risk factors for low birthweight. *Am J Obstet Gynecol* 2010; **202**: 572.e1–6.
116 Mamabolo RL, Alberts M, Steyn NP, Levitt NS. The effect of maternal glucose metabolism, iron, vitamin B12 and folate status on pregnancy outcomes. *South African J Clin Nutr* 2006; **19**: 120–30.
117 Allen LH, Rosenberg IH, Oakley GP, Omenn GS. Considering the case for vitamin B12 fortification of flour. *Food Nutr Bull* 2010; **31**: S36–46.
118 Forsdahl A. Are poor living conditions in childhood and adolescence an important risk factor for arteriosclerotic heart disease? *Br J Prev Soc Med* 1977; **31**: 91–5.
119 Barker D, Osmond C. Infant Mortality, Childhood Nutrition, and Ischaemic Heart Disease in England and Wales. *Lancet* 1986; **327**: 1077–81.
120 Hattersley AT, Tooke JE. The fetal insulin hypothesis: an alternative explanation of the association of low birthweight with diabetes and vascular disease. *Lancet* 1999; **353**: 1789–92.
121 Desai M, Jellyman JK, Ross MG. Epigenomics, gestational programming and risk of metabolic syndrome. *Int J Obes* 2015; **39**: 633–41.
122 Gluckman PD, Hanson MA, Cooper C, Thornburg KL. Effect of in utero and early-life conditions and adult health and disease. *N Engl J Med* 2008; **359**: 1523–4; author reply 1524.
123 Bock J, Wainstock T, Braun K, Segal M. Stress In Utero: Prenatal Programming of Brain Plasticity and Cognition. *Biol Psychiatry* 2015; **78**: 315–26.
124 Tau GZ, Peterson BS. Normal development of brain circuits. *Neuropsychopharmacology* 2010; **35**: 147–68.
125 Hüppi PS, Warfield S, Kikinis R, *et al.* Quantitative magnetic resonance imaging of brain development in premature and mature newborns. *Ann Neurol* 1998; **43**: 224–35.
126 Connors SL, Levitt P, Matthews SG, *et al.* Fetal mechanisms in neurodevelopmental disorders. *Pediatr Neurol* 2008; **38**: 163–76.
127 Krikke G, Grooten I, Vrijkotte T, van Eijsden M, Roseboom T, Painter R. Vitamin B12 and folate status in early pregnancy and cardiometabolic risk factors in the offspring at age 5-6 years: findings from the ABCD multi-ethnic birth cohort. *BJOG An Int J Obstet Gynaecol* 2016; **123**: 384–92.
128 Lovblad K, Ramelli G, Remonda L, Nirkko AC, Ozdoba C, Schroth G. Retardation of myelination due to dietary vitamin B12 deficiency: cranial MRI findings. *Pediatr Radiol* 1997; **111**: 155–8.

129 Peterson BS, Vohr B, Staib LH, *et al.* Regional brain volume abnormalities and long-term cognitive outcome in preterm infants. *JAMA* 2000; **284**: 1939–47.
130 Martinussen M, Flanders DW, Fischl B, *et al.* Segmental brain volumes and cognitive and perceptual correlates in 15-year-old adolescents with low birth weight. *J Pediatr* 2009; **155**: 848–53.e1.
131 Tolsa CB, Zimine S, Warfield SK, *et al.* Early alteration of structural and functional brain development in premature infants born with intrauterine growth restriction. *Pediatr Res* 2004; **56**: 132–8.
132 De Bie HMA, Oostrom KJ, Boersma M, *et al.* Global and regional differences in brain anatomy of young children born small for gestational age. *PLoS One* 2011; **6**: e24116.
133 Ardila a. A neuropsychological approach to intelligence. *Neuropsychol Rev* 1999; **9**: 117–36.
134 Breslau N, Paneth N, Lucia V, Paneth-Pollak R. Maternal smoking during pregnancy and offspring IQ. *Int J Epidemiol* 2005; **34**: 1047–53.
135 Østgård HF, Sølsnes AE, Bjuland KJ, *et al.* Executive function relates to surface area of frontal and temporal cortex in very-low-birth-weight late teenagers. *Early Hum Dev* 2016; **95**: 47–53.
136 Peterson BS, Anderson AW, Ehrenkranz R, *et al.* Regional brain volumes and their later neurodevelopmental correlates in term and preterm infants. *Pediatrics* 2003; **111**: 939–48.
137 Moster D, Lie RT, Markestad T. Long-term medical and social consequences of preterm birth. *N Engl J Med* 2008; **359**: 262–73.
138 Hollo O, Rautava P, Korhonen T, Helenius H, Kero P, Sillanpää M. Academic achievement of small-for-gestational-age children at age 10 years. *Arch Pediatr Adolesc Med* 2002; **156**: 179–87.
139 Strauss RS. Adult functional outcome of those born small for gestational age: twenty-six-year follow-up of the 1970 British Birth Cohort. *JAMA* 2000; **283**: 625–32.
140 Paz I, Gale R, Laor A, Danon YL, Stevenson DK, Seidman DS. The cognitive outcome of full-term small for gestational age infants at late adolescence. *Obstet Gynecol* 1995; **85**: 452–6.
141 Theodore RF, Thompson JMD, Waldie KE, *et al.* Determinants of cognitive ability at 7 years: a longitudinal case-control study of children born small-for-gestational age at term. *Eur J Pediatr* 2009; **168**: 1217–24.
142 Frangou S, Chitins X, Williams SCR. Mapping IQ and gray matter density in healthy young people. *Neuroimage* 2004; **23**: 800–5.
143 Allin M, Matsumoto H, Santhouse a M, *et al.* Cognitive and motor function and the size of the cerebellum in adolescents born very pre-term. *Brain* 2001; **124**: 60–6.
144 Mallard C, Loeliger M, Copolov D, Rees S. Reduced number of neurons in the hippocampus and the cerebellum in the postnatal guinea-pig following intrauterine growth-restriction. *Neuroscience* 2000; **100**: 327–33.
145 Allin MPG. Novel insights from quantitative imaging of the developing cerebellum. *Semin Fetal Neonatal Med* 2016; : 1–6.
146 Tramo MJ, Loftus WC, Stukel T a, Green RL, Weaver JB, Gazzaniga MS. Brain size, head size, and intelligence quotient in monozygotic twins. *Neurology* 1998; **50**: 1246–52.
147 University of York. PROSPERO. 2015. http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42015025141 (accessed Jan 15, 2016).
148 The Cochrane Collaboration. Cochrane Handbook for Systematic Reviews of Interventions, 5.1.0 edn. 2011.

149 Wells G, Shea B, O'Connell D, *et al.* The Newcastle-Ottawa Scale for assessing the quality of nonrandomised studies in meta-analyses. 2013. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.
150 Andreassen S. Cognitive Function in Small for Gestational Age (SGA) individuals born at term – from early childhood to young adulthood. 2008; : 1–35.
151 Skranes JS, Martinussen M, Smevik O, *et al.* Cerebral MRI findings in very-low-birth-weight and small-for-gestational-age children at 15 years of age. *Pediatr Radiol* 2005; **35**: 758–65.
152 Gjessing HK, Grøttum P, Eik-Nes SH. A direct method for ultrasound prediction of day of delivery: a new, population-based approach. *Ultrasound Obstet Gynecol* 2007; **30**: 19–27.
153 Gjessing H, Grøttum P, Økland I, Dreier J, Eik-Nes S. [eSnurra Growth - Description of Methods]. 2012; : 1–8.
154 Maslovitz S, Shimonovitz S, Lessing JB, Hochner-Celnikier D. The validity of oral glucose tolerance test after 36 weeks' gestation. *Eur J Obstet Gynecol Reprod Biol* 2006; **129**: 19–24.
155 Jørgensen LGM, Schytte T, Brandslund I, Stahl M, Petersen PH, Andersen B. Fasting and post-glucose load--reference limits for peripheral venous plasma glucose concentration in pregnant women. *Clin Chem Lab Med* 2003; **41**: 187–99.
156 Hebel JR, Fox NL, Sexton M. Dose-response of birth weight to various measures of maternal smoking during pregnancy. *J Clin Epidemiol* 1988; **41**: 483–9.
157 Wechsler D. Manual for the Wechsler Preschool and Primary Scale of Intelligence - Revised (WPPSI-R). San Antonio, TX: Psychological Corporation, 1989.
158 Wechsler D. Manual for the Wechsler Intelligence Scale for Children. New York: Psychological Corporation, 1991.
159 Martinussen M, Fischl B, Larsson HB, *et al.* Cerebral cortex thickness in 15-year-old adolescents with low birth weight measured by an automated MRI-based method. *Brain* 2005; **128**: 2588–96.
160 Fischl B, Salat DH, Busa E, *et al.* Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 2002; **33**: 341–55.
161 Jack CR, Twomey CK, Zinsmeister AR, Sharbrough FW, Petersen RC, Cascino GD. Anterior temporal lobes and hippocampal formations: normative volumetric measurements from MR images in young adults. *Radiology* 1989; **172**: 549–54.
162 Stabler SP. Vitamin B12 Deficiency. *N Engl J Med* 2013; **368**: 149–60.
163 Villar J, Ismail LC, Victora CG, *et al.* International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. *Lancet* 2014; **384**: 857–68.
164 Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 2001; **29**: 1165–88.
165 McNutt L-A. Estimating the Relative Risk in Cohort Studies and Clinical Trials of Common Outcomes. *Am J Epidemiol* 2003; **157**: 940–3.
166 Nurminen M. To use or not to use the odds ratio in epidemiologic analyses? *Eur J Epidemiol* 1995; **11**: 365–71.
167 Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol* 2004; **159**: 702–6.
168 Berlin I. Smoking-induced metabolic disorders: A review. *Diabetes Metab* 2008; **34**: 307–14.
169 Pollack H, Lantz PM, Frohna JG. Maternal smoking and adverse birth outcomes among singletons and twins. *Am J Public Health* 2000; **90**: 395–400.
170 Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment

in epidemiologic studies. *Epidemiology* 2009; **20**: 488–95.
171 Stewart GB, Altman DG, Askie LM, Duley L, Simmonds MC, Stewart L a. Statistical Analysis of Individual Participant Data Meta-Analyses: A Comparison of Methods and Recommendations for Practice. *PLoS One* 2012; **7**: 1–8.
172 Debray TP a, Moons KGM, Abo-Zaid GMA, Koffijberg H, Da Riley R. Individual Participant Data Meta-Analysis for a Binary Outcome: One-Stage or Two-Stage? *PLoS One* 2013; **8**: 1–10.
173 Lau J, Ioannidis JP, Schmid CH. Quantitative synthesis in systematic reviews. *Ann Intern Med* 1997; **127**: 820–6.
174 The World Bank. The World Bank. 2016. http://data.worldbank.org/country (accessed Jan 15, 2016).
175 Knight BA, Shields BM, Brook A, *et al.* Lower Circulating B12 Is Associated with Higher Obesity and Insulin Resistance during Pregnancy in a Non-Diabetic White British Population. *PLoS One* 2015; **10**: e0135268.
176 Krishnaveni G V., Hill JC, Veena SR, *et al.* Low plasma vitamin B12 in pregnancy is associated with gestational 'diabesity' and later diabetes. *Diabetologia* 2009; **52**: 2350–8.
177 Karmi O, Zayed A, Baraghethi S, Qadi M, Ghanem R. Measurement of vitamin B12 concentration: A review on available methods. *IIOAB J* 2011; **2**: 23–32.
178 Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**: 629–34.
179 Higgins JPT, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ Br Med J* 2003; **327**: 557–60.
180 Sterne J a C, White IR, Carlin JB, *et al.* Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ* 2009; **338**: b2393–b2393.
181 He Y. Missing data analysis using multiple imputation: getting to the heart of the matter. *Circ Cardiovasc Qual Outcomes* 2010; **3**: 98–105.
182 Horton NJ, Lipsitz SR, Parzen M. A Potential for Bias When Rounding in Multiple Imputation. *Am Stat* 2003; **57**: 229–32.
183 Koopman L, Van Der Heijden GJMG, Grobbee DE, Rovers MM. Comparison of methods of handling missing data in individual patient data meta-analyses: An empirical example on antibiotics in children with acute otitis media. *Am J Epidemiol* 2008; **167**: 540–5.
184 Jolani S, Debray TP a., Koffijberg H, van Buuren S, Moons KGM. Imputation of systematically missing predictors in an individual participant data meta-analysis: a generalized approach using MICE. *Stat Med* 2015; **34**: 1841–63.
185 Rothman KJ, Greenland S, Lash TL. Modern Epidemiology, 3rd edn. Lippincott Williams & Wilkins, 2008.
186 Hill AB. The Environment and Disease: Association or Causation? *Proc R Soc Med* 1965; **58**: 295–300.
187 Rothman KJ, Greenland S. Causation and causal inference in epidemiology. *Am J Public Health* 2005; **95**: 144–50.
188 Mooy JM, Grootenhuis PA, de Vries H, *et al.* Intra-individual variation of glucose, specific insulin and proinsulin concentrations measured by two oral glucose tolerance tests in a general Caucasian population: the Hoorn Study. *Diabetologia* 1996; **39**: 298–305.
189 Ma KK, Mele L, Landon MB, *et al.* The obstetric and neonatal implications of a low value on the 50-g glucose screening test. *Am J Perinatol* 2013; **30**: 715–22.
190 Kline RB, Snyder J, Guilmette S, Castellanos M. External validity of the profile

variability index for the K-ABC, Stanford-Binet, and WISC-R: another cul de-sac. *J Learn Disabil* 1993; **26**: 557–67.
191 Chen LW, Lim AL, Colega M, *et al.* Maternal folate status, but not that of vitamins B-12 or B-6, is associated with gestational age and preterm birth risk in a multiethnic Asian population. *J Nutr* 2015; **145**: 113–20.
192 Rothman KJ. Epidemiology - An introduction, 1st edn. New York: Oxford University Press, 2002.
193 Folkehelseinstituttet. Årstabeller for Medisinsk fødselsregister 2011 – Fødsler i Norge. 2013.
194 Shinohara S, Hirai M, Hirata S, Suzuki K. Relation between low 50-g glucose challenge test results and small-for-gestational-age infants. *J Obstet Gynaecol Res* 2015; **41**: 1752–6.
195 Yajnik CS, Chandak GR, Joglekar C, *et al.* Maternal homocysteine in pregnancy and offspring birthweight: epidemiological associations and Mendelian randomization analysis. *Int J Epidemiol* 2014; **43**: 1487–97.
196 Langer O, Levy J, Brustman L, Anyaegbunam A, Merkatz R, Divon M. Glycemic control in gestational diabetes mellitus--how tight is tight enough: small for gestational age versus large for gestational age? *Am J Obstet Gynecol* 1989; **161**: 646–53.
197 Bhat R, Bhagya K, Kumar P. Association of Low Maternal Plasma Glucose after Oral Glucose Challenge Test with Small for Gestational Age Neonate. *Int J Infertil Fetal Med* 2012; **3**: 22–5.
198 Cox DR. Journal of the American Note on Grouping. *J Am Stat Assoc* 1957; **52**: 543–7.
199 Bernstein I. Fetal body composition. *Curr Opin Clin Nutr Metab Care* 2005; **8**: 613–7.
200 Abell D a., Beischer N a. The Relationship Between Maternal Glucose Tolerance and Fetal Size at Birth. *Aust N Z J Obs Gynaecol* 1976; **16**: 1–4.
201 Langer O, Damus K, Maiman M, Divon M, Levy J, Bauman W. A link between relative hypoglycemia-hypoinsulinemia during oral glucose tolerance tests and intrauterine growth retardation. *Am J Obs Gynecol* 1986; **155**: 711–6.
202 Vadakekut ES, McCoy SJB, Payton ME. Association of maternal hypoglycemia with low birth weight and low placental weight: a retrospective investigation. *J Am Osteopath Assoc* 2011; **111**: 148–52.
203 Bienstock JL, Holcroft CJ, Althaus J. Small fetal abdominal circumference in the second trimester and subsequent low maternal plasma glucose after a glucose challenge test is associated with the delivery of a small-for-gestational age neonate. *Ultrasound Obstet Gynecol* 2008; **31**: 517–9.
204 Weissman A, Solt I, Zloczower M, Jakobi P. Hypoglycemia during the 100-g oral glucose tolerance test: incidence and perinatal significance. *Obstet Gynecol* 2005; **105**: 1424–8.
205 Sokol RJ, Kazzi GM, Kalhan SC, Pillay SK. Identifying the pregnancy at risk for intrauterine growth retardation: possible usefulness of the intravenous glucose tolerance test. *Am J Obs Gynecol* 1982; **143**: 220–3.
206 Khouzami VA, Ginsburg DS, Daikoku NH, Johnson JW. The glucose tolerance test as a means of identifying intrauterine growth retardation. *Am J Obs Gynecol* 1981; **139**: 423–6.
207 Piper JM, Field NT, Higby K, Elliott BD, Langer O. Maternal-fetal glucose metabolism and fetal growth retardation. Is there an association? *J Reprod Med* 1996; **41**: 761–6.
208 Caruso A, Paradisi G, Ferrazzani S, Lucchese A, Moretti S, Fulghesu AM. Effect of maternal carbohydrate metabolism on fetal growth. *Obstet Gynecol* 1998; **92**: 8–12.
209 Holm BC, Svensson J, Åkesson C, *et al.* Evidence for immunological priming and increased frequency of CD4 + CD25+ cord blood T cells in children born to mothers

with type 1 diabetes. *Clin Exp Immunol* 2006; **146**: 493–502.
210 Bergen NE, Jaddoe VW V, Timmermans S, *et al.* Homocysteine and folate concentrations in early pregnancy and the risk of adverse pregnancy outcomes: The Generation R Study. *BJOG* 2012; **119**: 739–51.
211 Dhobale M, Chavan P, Kulkarni A, Mehendale S, Pisal H, Joshi S. Reduced folate, increased vitamin B(12) and homocysteine concentrations in women delivering preterm. *Ann Nutr Metab* 2012; **61**: 7–14.
212 Ronnenberg AG, Goldman MB, Chen D, *et al.* Preconception homocysteine and B vitamin status and birth outcomes in Chinese women. *Am J Clin Nutr* 2002; **76**: 1385–91.
213 Miner SES, Evrovski J, Cole DEC. Clinical chemistry and molecular biology of homocysteine metabolism: An update. *Clin Biochem* 1997; **30**: 189–201.
214 Lindblad B, Zaman S, Malik A, *et al.* Folate, vitamin B12, and homocysteine levels in South Asian women with growth-retarded fetuses. *Acta Obstet Gynecol Scand* 2005; **84**: 1055–61.
215 Sonne S, Bhalla V, Barman S, *et al.* Hyperhomocysteinemia is detrimental to pregnancy in mice and is associated with preterm birth. *Biochim Biophys Acta - Mol Basis Dis* 2013; **1832**: 1149–58.
216 Wadhwani NS, Patil V V, Mehendale SS, Wagh GN, Gupte SA, Joshi SR. Increased homocysteine levels exist in women with preeclampsia from early pregnancy. *J Matern Fetal Neonatal Med* 2016; **29**: 2719–25.
217 Ge J, Wang J, Zhang F, *et al.* Correlation between MTHFR gene methylation and pre-eclampsia, and its clinical significance. *Genet Mol Res* 2015; **14**: 8021–8.
218 Makedos G, Papanicolaou A, Hitoglou A, *et al.* Homocysteine, folic acid and B12 serum levels in pregnancy complicated with preeclampsia. *Arch Gynecol Obstet* 2007; **275**: 121–4.
219 Mignini LE, Latthe PM, Villar J, Kilby MD, Carroli G, Khan KS. Mapping the theories of preeclampsia: the role of homocysteine. *Obstet Gynecol* 2005; **105**: 411–25.
220 Scholl TO, Johnson WG. Folic acid: influence on the outcome of pregnancy. *Am J Clin Nutr* 2000; **71**: 1295S – 303S.
221 Martinussen MP, Bracken MB, Triche EW, Jacobsen GW, Risnes KR. Folic acid supplementation in early pregnancy and the risk of preeclampsia, small for gestational age offspring and preterm delivery. *Eur J Obstet Gynecol Reprod Biol* 2015; **195**: 94–9.
222 Knudtson EJ, Smith K, Mercer BM, *et al.* Serum homocysteine levels after preterm premature rupture of the membranes. *Am J Obstet Gynecol* 2004; **191**: 537–41.
223 Refsum H, Ueland PM, Nygård O, Vollset SE. Homocysteine and cardiovascular disease. *Annu Rev Med* 1998; **49**: 31–62.
224 Hogeveen M, Blom HJ, den Heijer M. Maternal homocysteine and small-for-gestational-age offspring: systematic review and meta-analysis. *Am J Clin Nutr* 2012; **95**: 130–6.
225 Lassi ZS, Salam RA, Haider BA, Bhutta ZA. Folic acid supplementation during pregnancy for maternal health and pregnancy outcomes. *Cochrane database Syst Rev* 2013; **3**: CD006896.
226 Duggan C, Srinivasan K, Thomas T, *et al.* Vitamin B-12 supplementation during pregnancy and early lactation increases maternal, breast milk, and infant measures of vitamin B-12 status. *J Nutr* 2014; **144**: 758–64.
227 McCullough LE, Miller EE, Mendez MA, Murtha AP, Murphy SK, Hoyo C. Maternal B vitamins: effects on offspring weight and DNA methylation at genomically imprinted domains. *Clin Epigenetics* 2016; **8**: 8.
228 Sinclair KD, Allegrucci C, Singh R, *et al.* DNA methylation, insulin resistance, and blood pressure in offspring determined by maternal periconceptional B vitamin and

methionine status. *Proc Natl Acad Sci U S A* 2007; **104**: 19351–6.
229 Yajnik CS, Deshpande SS, Jackson a a, *et al.* Vitamin B12 and folate concentrations during pregnancy and insulin resistance in the offspring: the Pune Maternal Nutrition Study. *Diabetologia* 2008; **51**: 29–38.
230 Liu Q, Li S, Quan H, Li J. Vitamin B12 status in metformin treated patients: Systematic review. *PLoS One* 2014; **9**: 8–13.
231 Molloy AM, Kirke PN, Brody LC, Scott JM, Mills JL. Effects of folate and vitamin B12 deficiencies during pregnancy on fetal, infant, and child development. *Food Nutr Bull* 2008; **29**: S101–11; discussion S112–5.
232 Tanis JC, Van Braeckel KNJ a., Kerstjens JM, Bocca-Tjeertes IF a., Reijneveld S a., Bos AF. Functional Outcomes at Age 7 Years of Moderate Preterm and Full Term Children Born Small for Gestational Age. *J Pediatr* 2015; **166**: 552–8.e1.
233 Sommerfelt K. Cognitive development of term small for gestational age children at five years of age. *Arch Dis Child* 2000.
234 Chen J, Chen P, Bo T, Luo K. Cognitive and Behavioral Outcomes of Intrauterine Growth Restriction School-Age Children. *Pediatrics* 2016; **137**. DOI:10.1542/peds.2015-3868.
235 Jensen RB, Juul A, Larsen T, Mortensen EL, Greisen G. Cognitive ability in adolescents born small for gestational age: Associations with fetal growth velocity, head circumference and postnatal growth. *Early Hum Dev* 2015; **91**: 755–60.
236 Andersson HW, Sommerfelt K, Sonnander K, Ahlsten G. Maternal child-rearing attitudes, IQ, and socioeconomic status as related to cognitive abilities of five-year-old children. *Psychol Rep* 1996; **79**: 3–14.
237 Lange N, Froimowitz MP, Bigler ED, Lainhart JE. Associations between IQ, total and regional brain volumes, and demography in a large normative sample of healthy children and adolescents. *Dev Neuropsychol* 2010; **35**: 296–317.
238 Resnik R. Intrauterine growth restriction. *Obstet Gynecol* 2002; **99**: 490–6.
239 Mallard EC, Rees S, Stringer M, Cock ML, Harding R. Effects of Chronic Placental Insufficiency on Brain Development in Fetal Sheep. *Pediatr Res* 1998; **43**: 262–70.
240 Rehn a E, Van Den Buuse M, Copolov D, Briscoe T, Lambert G, Rees S. An animal model of chronic placental insufficiency: relevance to neurodevelopmental disorders including schizophrenia. *Neuroscience* 2004; **129**: 381–91.
241 Vangberg TR, Skranes J, Dale AM, Martinussen M, Brubakk A-M, Haraldseth O. Changes in white matter diffusion anisotropy in adolescents born prematurely. *Neuroimage* 2006; **32**: 1538–48.
242 Pierson CR, Folkerth RD, Billiards SS, *et al.* Gray matter injury associated with periventricular leukomalacia in the premature infant. *Acta Neuropathol* 2007; **114**: 619–31.
243 Volpe JJ. Brain injury in premature infants: a complex amalgam of destructive and developmental disturbances. *Lancet Neurol* 2009; **8**: 110–24.
244 Ligam P, Haynes RL, Folkerth RD, *et al.* Thalamic damage in periventricular leukomalacia: novel pathologic observations relevant to cognitive deficits in survivors of prematurity. *Pediatr Res* 2009; **65**: 524–9.
245 Boardman JP, Counsell SJ, Rueckert D, *et al.* Abnormal deep grey matter development following preterm birth detected using deformation-based morphometry. *Neuroimage* 2006; **32**: 70–8.
246 Marin-Padilla M. Developmental Neuropathology and Impact of Perinatal Brain Damage. III. *J Neuropathol Exp Neurol* 1999; **58**: 407–29.
247 Chang CH, Chang FM, Yu CH, Ko HC, Chen HY. Assessment of fetal cerebellar volume using three-dimensional ultrasound. *Ultrasound Med Biol* 2000; **26**: 981–8.

248 Takahashi E, Hayashi E, Folkerth RD, Schmahmann JD, Ellen Grant P. Development of cerebellar connectivity in human fetal brains revealed by high angular resolution diffusion tractography. *Neuroimage* 2014; **96**: 326–33.
249 Limperopoulos C, Soul JS, Haidar H, *et al.* Impaired trophic interactions between the cerebellum and the cerebrum among preterm infants. *Pediatrics* 2005; **116**: 844–50.
250 Pierson CR, Al Sufiani F. Preterm birth and cerebellar neuropathology. *Semin Fetal Neonatal Med* 2016; : 1–7.
251 Letinic K, Rakic P. Telencephalic origin of human thalamic GABAergic neurons. *Nat Neurosci* 2001; **4**: 931–6.
252 Shaw P, Greenstein D, Lerch J, *et al.* Intellectual ability and cortical development in children and adolescents. *Nature* 2006; **440**: 676–9.
253 Schnack HG, van Haren NEM, Brouwer RM, *et al.* Changes in Thickness and Surface Area of the Human Cortex and Their Relationship with Intelligence. *Cereb Cortex* 2014; **10**. DOI:10.1093/cercor/bht357.
254 Simões R V., Cruz-Lemini M, Bargalló N, Gratacós E, Sanz-Cortes M. Brain Metabolite Differences in One-Year-Old Infants Born Small at Term and its Association with Neurodevelopmental Outcome. *Am J Obstet Gynecol* 2015; **213**: 1–11.
255 Nasrat H. Use of ultrasound longitudinal data in the diagnosis of abnormal fetal growth. *J Matern Fetal Med* 1997; **6**: 209–14.
256 Baumann MU, Deborde S, Illsley NP. Placental glucose transfer and fetal growth. *Endocrine* 2002; **19**: 13–22.
257 Brodszki J, Marsal K. Management of pregnancies with suspected intrauterine growth retardation in Sweden. *Acta Obstet Gynecol Scand* 2000; **79**: 723–8.
258 Løhaugen GCC, Østgård HF, Andreassen S, *et al.* Small for Gestational Age and Intrauterine Growth Restriction Decreases Cognitive Function in Young Adults. *J Pediatr* 2013; **163**: 447–53.
259 Walker D-M, Marlow N. Neurocognitive outcome following fetal growth restriction. *Arch Dis Child Fetal Neonatal Ed* 2008; **93**: F322–5.
260 Stoknes M, Andersen GL, Dahlseng MO, *et al.* Cerebral palsy and neonatal death in term singletons born small for gestational age. *Pediatrics* 2012; **130**: e1629–35.
261 Topp M, Langhoff-Roos J, Uldall P, Kristensen J. Intrauterine growth and gestational age in preterm infants with cerebral palsy. *Early Hum Dev* 1996; **44**: 27–36.
262 Dwarkanath P, Barzilay JR, Thomas T, Thomas A, Bhat S, Kurpad A. High folate and low vitamin B-12 intakes during pregnancy are associated with small-for-gestational age infants in South Indian women: a prospective observational cohort study. *Am J Clin Nutr* 2013; **98**: 1450–8.
263 Blettner M, Sauerbrei W, Schlehofer B, Scheuchenpflug T, Friedenreich C. Traditional reviews, meta-analyses and pooled analyses in epidemiology. *Int J Epidemiol* 1999; **28**: 1–9.
264 Position of the American Dietetic Association and Dietitians of Canada: Vegetarian diets. *J Am Diet Assoc* 2003; **103**: 748–65.
265 Ota E, Mori R, Middleton P, *et al.* Zinc supplementation for improving pregnancy and infant outcome. *Cochrane database Syst Rev* 2015; **2**: CD000230.
266 De-Regil LM, Palacios C, Lombardo LK, Peña-Rosas JP. Vitamin D supplementation for women during pregnancy. *Cochrane database Syst Rev* 2016; **1**: CD008873.
267 Wald DS, Law M, Morris JK. Homocysteine and cardiovascular disease: evidence on causality from a meta-analysis. *Bmj* 2002; **325**: 1202.
268 Bønaa KH, Njølstad I, Ueland PM, *et al.* Homocysteine lowering and cardiovascular events after acute myocardial infarction. *N Engl J Med* 2006; **354**: 1578–88.
269 Shenkin SD, Starr JM, Deary IJ. Birth weight and cognitive ability in childhood: a

systematic review. *Psychol Bull* 2004; **130**: 989–1013.
270 Arcangeli T, Thilaganathan B, Hooper R, Khan KS, Bhide A. Neurodevelopmental delay in small babies at term: a systematic review. *Ultrasound Obstet Gynecol* 2012; **40**: 267–75.
271 Spittle A, Orton J, Anderson PJ, Boyd R, Doyle LW. Early developmental intervention programmes provided post hospital discharge to prevent motor and cognitive impairment in preterm infants. *Cochrane database Syst Rev* 2015; **11**: CD005495.

Appendix

- Paper 1
    - Main document, including tables and figures
    - Supplementary tables
- Paper 2
    - Main document, including tables and figures
    - Supplementary text and tables
- Paper 3
    - Main document
    - Tables to main document
    - Figures to main document
    - Supplementary text, tables and figures

Paper I

AOGS
ACTA Obstetricia et Gynecologica Scandinavica

AOGS MAIN RESEARCH ARTICLE

Association between low blood glucose increase during glucose tolerance tests in pregnancy and impaired fetal growth

TORMOD ROGNE & GEIR W. JACOBSEN

Department of Public Health and General Practice, Faculty of Medicine, Norwegian University of Science and Technology, Trondheim, Norway

Key words
Fetal development, fetal growth retardation, glucose tolerance test, blood glucose, glucose metabolism disorders

Correspondence
Tormod Rogne, Department of Public Health and General Practice, Faculty of Medicine, Norwegian University of Science and Technology, Postbox 8905, 7491 Trondheim, Norway.
E-mail: tormorog@stud.ntnu.no

Conflicts of interest
The authors have stated explicitly that there are no conflicts of interest in connection with this article, nor financial relationships relevant to this article to disclose.

Please cite this article as: Rogne T, Jacobsen GW. Association between low blood glucose increase during glucose tolerance tests in pregnancy and impaired fetal growth. Acta Obstet Gynecol Scand 2014; 93: 1160–1169.

Received: 3 July 2013
Accepted: 17 February 2014

DOI: 10.1111/aogs.12365

Abstract

Objective. To evaluate how different levels of increase in maternal blood glucose from a fasting state to 2 h after an oral glucose challenge in late pregnancy are associated with fetal growth, with special emphasis on those with a low increase. *Design.* Prospective cohort study. *Subjects.* We followed 855 women, of whom 70% had an increased risk for carrying lighter babies. *Study design and methods.* Ultrasound was used to estimate fetal growth in gestational weeks 25, 33 and 37. In week 37 the women had a 75-g oral glucose tolerance test, and fasting and 2-h capillary glucose values were recorded with the difference between these two called delta (Δ) glucose. Three groups were constructed from the Δ glucose distribution: Low below the 10th centile; Medium between the 10th and 90th centiles; and High above the 90th centile. Missing data were imputed. Linear and Poisson regression models were applied. *Outcome measures.* Estimated fetal weight, percent deviation from expected fetal weight and anthropometric measures at birth. *Results.* The Low group carried the lightest fetuses and the High group the heaviest. The fetal growth in the Low group deviated increasingly more in a negative direction from week 25 to 37 than in the other groups. *Conclusion.* In a high-risk population, a positive relation between Δ glucose and fetal growth was found. The Low group demonstrated impaired growth. More attention should be paid to pregnant women with an insufficient increase in glucose after a glucose challenge. Future studies should challenge our findings in high-risk and low-risk populations.

Abbreviations: DM, overt diabetes mellitus; GDM, gestational diabetes mellitus; H, high Δ glucose; IUGR, intrauterine growth restriction; L, low Δ glucose; M, medium Δ glucose; OGTT, oral glucose tolerance test; SGA, small-for-gestational age.

Introduction

The intrauterine period is a defining time for an individual's life. Problems from psychiatric and behavioral disorders in childhood to cardiovascular diseases in the elderly have been associated with the intrauterine environment (1,2). An important assessment of fetal

Key Message

Excessive maternal insulin sensitivity in late pregnancy has been reported to negatively affect fetal growth. We found that an insufficient increase in maternal blood glucose after an oral glucose challenge in late pregnancy was associated with impaired fetal growth.

well-being is the pattern of growth. Identification of pregnancies with increased risk for giving birth to growth-restricted babies is critical for preventive and management strategies.

For fetal growth, changes in the maternal metabolism and distribution of glucose during pregnancy are of great importance. With increasing gestational age there is a corresponding increase in maternal insulin resistance (3). A consequence of this diabetogenic adaptation is an increase in maternal glucose concentration. Glucose is the most important nutrient for fetal growth and the main determinant for fetal glucose concentration is that of its mother (4). Insulin does not cross the placenta, so the fetus is solely dependent on its own insulin production, of which glucose is the strongest stimulant (5). Insulin is a key growth factor in fetal life, both directly and through stimulation of the insulin-like growth factor axis (6). The increased maternal insulin resistance in pregnancy therefore aids the fetus in two important ways: by increased supply of glucose (energy) and by stimulation of fetal insulin production (growth).

Excessive insulin resistance has been thoroughly investigated. It is characterized by high levels of maternal and fetal glucose and insulin, and may lead to gestational diabetes mellitus (GDM) or overt diabetes mellitus (DM) (5,7). Fetal overgrowth may be one of several short-term consequences (8). Far fewer studies have been conducted on women at the other end of the insulin resistance spectrum, i.e. those who are too insulin sensitive. Still, that group may be as important, because there is reason to suspect impaired fetal growth in pregnant woman with a suboptimal increase in insulin resistance (9).

There are several good measures for insulin resistance, but most of them are complex (10). The oral glucose tolerance test (OGTT) is a simple and widely applied test that evaluates glucose tolerance from recordings of fasting and 2-h glucose values. We assumed that the difference between the 2-h and fasting values, termed Δ glucose, may serve as a proxy for glucose tolerance.

Much is still unknown regarding the causes of intrauterine growth restriction (IUGR). The aim of this study was to evaluate the association between different levels of Δ glucose and fetal growth in a pregnant population with increased risk of IUGR, with special emphasis on women with low Δ glucose. Our high-risk population was deemed appropriate to explore this aim as a high proportion of negative events would increase the power of the study. The main hypothesis was that a low Δ glucose is associated with restricted growth. Conversely, high Δ glucose was assumed to be strongly correlated with GDM and DM, and we expected an accelerated fetal growth in these pregnancies.

Material and methods

The data have been derived from a large multicenter population-based cohort known as the Scandinavian small-for-gestational age (SGA) Study (11). The overall aim was to study fetal growth, perinatal outcome and the tendency to repeat a negative outcome in consecutive births. Data were prospectively collected between January 1986 and March 1988 at the university hospitals of Trondheim and Bergen (Norway), and Uppsala (Sweden).

Figure 1 shows the selection of the present study population. Only para 1 and 2 women carrying singletons were followed. Of the original 1945 women eligible for detailed follow up (study visits at gestational weeks 17, 25, 33 and 37, and at birth), 860 (44%) completed an OGTT in week 37. This was done according to the protocol and regardless of any risk factors for GDM (11). Pregnancies where an OGTT was not performed were excluded, as were women with pre-pregnancy DM (n = 3). Since the OGTT took place in gestational week 37, only two of the women in our population had a preterm delivery. To avoid mixing the effects of preterm birth and fetal growth restriction, both pregnancies were excluded, which left us with a final study population of 855 pregnant women. About 70% of the women were characterized with some degree of increased risk for giving birth to an SGA baby while the remaining 30% were derived from a 10% random sample and served as a population reference (11).

All participants were informed about the study aims and objectives and gave their written consent. Both the Norwegian Data Inspectorate and the Regional Committee for Medical and Health Research Ethics approved the study. Further details of the background study have previously been reported (11).

By use of *eSnurra*, a well-established mathematical model developed by the National Center for Fetal Medicine, St Olavs University Hospital and Norwegian University of Science and Technology, Trondheim (12), gestational age and expected date of delivery were estimated from ultrasound measurements of biparietal diameter at the first visit (17 weeks). Fetal weights at 25, 33 and 37 weeks of gestation were also estimated from *eSnurra*, based on biparietal diameter and mean abdominal diameter. Femur length was used when biparietal diameter was missing (n = 19). For recordings before day 167, weight deviations (not estimated weights) were ascertained by the use of the manual *eSnurra* plastic wheel (n = 14) (13). Weight deviation was expressed as% deviation from a population-based reference at each gestational age (12). Growth between gestational ages was calculated from the estimated growth deviations. Based on growth in the 10% random sample, IUGR was defined as fetal

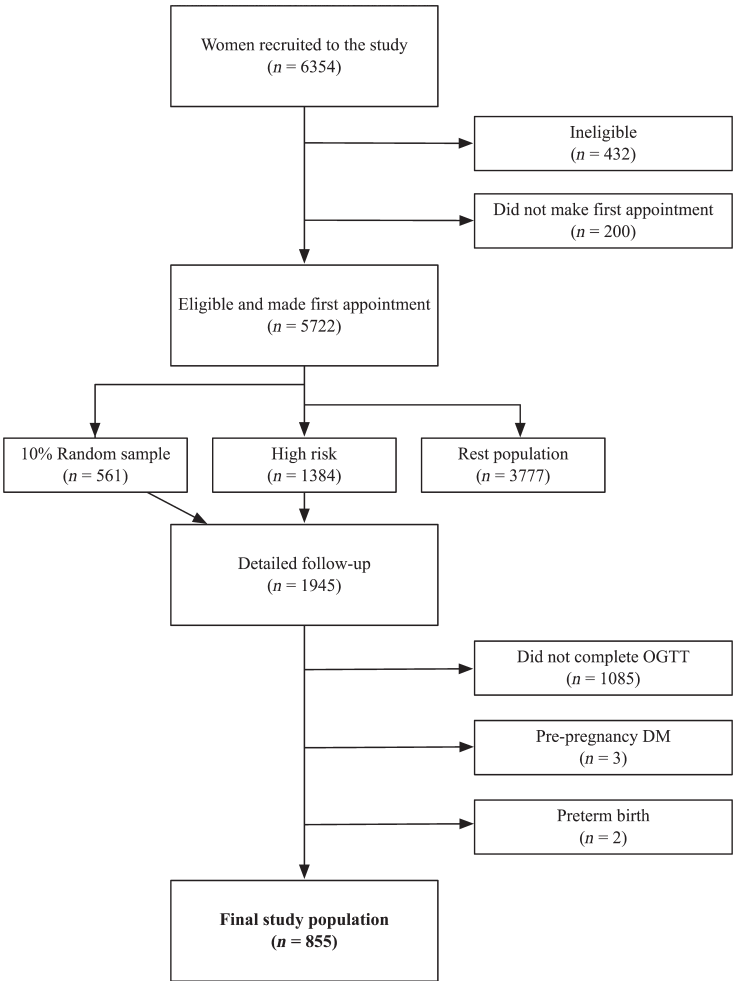


Figure 1. Flow chart for the selection of study participants, adapted from Bakketeig et al. (11). High-risk criteria (at least one): (i) previous low birth child, (ii) previous perinatal death, (iii) maternal low (<50 kg) pre-pregnancy weight, (iv) smoking around the time of conception, and (v) chronic maternal hypertension or renal disease. OGTT, oral glucose tolerance test; DM, overt diabetes mellitus.

growth from week 25 to week 37 below the 10th centile. SGA was defined as birthweight <10th centile at each gestational week, with reference standards specific for parity and newborn gender (11,14).

Capillary glucose was recorded after an overnight fast. Following consumption of 75 g glucose in 200 mL water over 10 min, the equivalent 2-h value was recorded (7,15). We defined Δ glucose (mmol/L) as 2-h glucose minus fasting glucose. From a 10% random sample we derived cut-offs for the 10th and the 90th percentiles of the Δ glucose distribution, which were 0.8 and 3.6 mmol/L, respectively. Three groups were therefore identified: Low group (L; Δ glucose ≤ 0.8 mmol/L; n = 103; 12%), High group (H; Δ glucose ≥ 3.6 mmol/L; n = 91; 11%), and Medium group (M; 0.8 mmol/L < Δ glucose < 3.6 mmol/L; n = 661; 77%).

The clinically familiar glucose metabolism disorders were defined as (capillary; mmol/L): hypoglycemia, fasting

and 2-h ≤2.80; impaired fasting glycemia, fasting ≥5.6 and <6.1, and 2-h <7.8; for GDM, fasting <6.1 and 2-h ≥7.8 and <11.1; and DM, fasting ≥6.1, 2-h ≥11.1 (7,15).

We considered the following established and potential confounders in the multivariate analyses: previous low birthweight child, mother's highest completed education, maternal age, parity, pre-pregnancy body mass index, weight gain, smoking habits during pregnancy, gestational age when the OGTT was performed, sex and gestational age at birth (11). Classification of smoking was based on reported number of cigarettes in week 33, after which the adverse effect on fetal growth seems greatest (16). In Table 1 levels of education and smoking are shown.

An influence-analysis of the fetal growth and growth outcome restricted to the 10% random sample was constructed to investigate how generalizable the findings of the main analysis were to a parous population with average risk.

Statistical analysis

All analyses were performed with the statistical program Stata 12.1 (Stata Corporation, College Station, TX, USA).

Normally distributed variables are presented as mean (95% CI), non-normally distributed variables as median (25th–75th centile), and frequencies as *n* (%). By use of generalized linear models, we fitted linear regression models for continuous outcome variables and Poisson regression models for binary outcome variables (17). For non-normally distributed outcome variables, we carried out the analyses with the appropriate nonidentity link function. Linear trend was tested by fitting regression models with the covariate representing the ordered Δ glucose groups as a continuous variable. Both univariate and multivariate (with potential confounders included) models were built. Robust standard errors were used in all models. Non-normally distributed continuous independent variables were transformed to the appropriate scale before analysis. We did not find interaction terms required.

Missing data were imputed (see the Supporting Information for details). For comparison, univariate and complete case analyses of all regression analyses may be found in Tables S1 and S2 (see the Supporting Information), with number of missing observations in the far right columns.

Results

Maternal baseline characteristics are reported in Table 1. Most characteristics were comparable between the groups. Pre-pregnancy weight tended towards higher weight with lower Δ glucose. Cigarette consumption was lower in group M than in the other two. Based on current Norwegian guidelines at least 81 (9%) of the women fulfilled the criteria for taking an OGTT as screening for GDM or DM (18).

By design, group L had the lowest Δ glucose and 2-h blood glucose values, and group H had the highest (Table 2). Intuitively, women in group L also had a slightly higher median fasting blood glucose, which helps to explain the five (5%) cases of DM in this group.

Table 1. Maternal baseline characteristics by Δ glucose group.

	Low (*n* = 103)	Medium (*n* = 661)	High (*n* = 91)
Height (cm), mean (95% CI)[a]	167.6 (166.5–168.8)	166.2 (165.8–166.7)	163.9 (162.7–165.1)
Age (years)[b]	28 (25–31)	28 (26–31)	28 (26–31)
Pre-pregnancy weight (kg)[b]	59 (54–68)	58 (52–63)	55 (48–60)
Pre-pregnancy body mass index (kg/m^2)[b]	21 (19–24)	21 (19–22)	20 (18–23)
Weight gain during pregnancy (kg)[b]	14 (11–17)	14 (11–17)	14 (11–17)
Years of education[c]			
≤9	20 (19)	108 (16)	21 (23)
10–12	56 (54)	378 (58)	44 (48)
≥13	27 (26)	171 (26)	26 (29)
Previous low birthweight child[c]	25 (24)	170 (26)	16 (18)
Para 1[c]	79 (77)	473 (72)	61 (67)
Number of cigarettes, week 33[c]			
0	37 (36)	319 (49)	40 (44)
1–9	26 (25)	144 (22)	19 (21)
10–19	34 (33)	159 (24)	28 (31)
≥20	5 (5)	32 (5)	4 (4)

Normally distributed variables presented as [a]mean (95% CI), [b]non-normally distributed variables presented as median (25th–75th centile) and [c]frequencies as *n* (%).

Table 2. Glucose profile by Δ glucose group.

	Low (n = 103)	Medium (n = 661)	High (n = 91)
Glycosylated hemoglobin week 37, %[a]	4.8 (4.7–5.0)	4.9 (4.8–4.9)	5.1 (5.0–5.2)
Glycosylated hemoglobin at delivery, %[a]	4.8 (4.7–4.9)	4.9 (4.9–4.9)	5.0 (4.9–5.1)
Fasting blood glucose week 37, mmol/L[b]	4.7 (4.4–5.1)	4.5 (4.2–4.9)	4.5 (4.1–5.0)
Two-hour blood glucose week 37, mmol/L[b]	5.0 (4.6–5.6)	6.6 (6.0–7.2)	9.0 (8.4–9.9)
Δ glucose week 37, mmol/L[b]	0.4 (0.0–0.6)	2.0 (1.5–2.6)	4.4 (4.0–5.0)
Gestational diabetes mellitus prior to week 37	0 (0)	3 (0)	0 (0)
Hypoglycemia week 37[c]	0 (0)	0 (0)	1 (1)
Impaired fasting glycemia week 37[c]	7 (7)	10 (2)	0 (0)
Gestational diabetes mellitus week 37[c]	0 (0)	66 (10)	78 (86)
Overt diabetes mellitus week 37[c]	5 (5)	6 (1)	9 (10)

Normally distributed variables presented as [a]mean (95% CI), [b]non-normally distributed variables presented as median (25th–75th centile) and [c]frequencies as n (%).

Eighty-seven (96%) of the group H women fulfilled the criteria for either GDM (n = 78; 86%) or DM (n = 9; 10%). The only mother with hypoglycemia was in group H. Glycosylated hemoglobin both in week 37 and at delivery support the overall impression that group H had the highest glucose load and group L the lowest.

Tables 3 and 4 present fetal growth pattern and growth outcome, respectively. The growth pattern of each group can be seen from Figure 2. Mean growth deviations in all three groups were comparable in week 25, although group H had a slightly higher estimated weight than the other two groups (−2.2%, −2.3% and −1.3% in groups L, M and H, respectively). In week 37 the differences between the three groups were most pronounced and group L deviated −5.1% from expected weight. Group L deviated by −2.4 percentage points more than group M. The linear trend analyses suggested a continuous relation between Δ glucose and fetal growth.

By comparing the weight deviation in week 37 with weight deviation at baseline (week 25) we obtained a measure of deviation from the group centiles. Both groups M and H followed their respective centiles and had more or less the same weight deviation in week 37 as

Table 3. Fetal growth by Δ glucose group; mean, comparison between groups and linear trend test (Low–Medium–High).

	Mean (95% CI)			Mean difference (95% CI), p-value				Mean (95% CI), p-value
	Low (n = 103)	Medium (n = 661)	High (n = 91)	Low vs. Medium		High vs Medium		Linear trend test
Estimated weight week 25, g[a]	781.7 (772.8–790.7)	776.6 (772.3–781.0)	788.1 (779.6, 796.6)	5.1 (−4.6 to 14.8)	0.31	11.4 (1.7–21.1)	0.02	2.9 (−3.5 to 9.3)	0.37
Estimated weight week 33, g	2019.8 (1990.8–2048.7)	2033.9 (2021.5–2046.3)	2073.3 (2043.5–2103.0)	−14.1 (−45.7 to 17.5)	0.38	39.4 (7.2–71.6)	0.02	26.1 (5.4–46.8)	0.01
Estimated weight week 37, g	2920.1 (2873.4–2966.7)	2993.7 (2973.8–3013.7)	3062.1 (3007.1–3117.1)	−73.6 (−124.5 to −22.8)	0.005	68.4 (10.1–126.7)	0.02	71.2 (35.6–106.8)	≤ 0.001
Growth deviation week 25, %	−2.2 (−3.1 to −1.2)	−2.3 (−2.7 to −1.9)	−1.3 (−2.1 to −0.4)	0.1 (−0.9 to 1.1)	0.87	1.0 (0.0–2.0)	0.04	0.4 (−0.2 to 1.1)	0.20
Growth deviation week 33, %	−4.4 (−5.8 to −3.1)	−3.8 (−4.4 to −3.2)	−1.8 (−3.2 to −0.4)	−0.6 (−2.1 to 0.8)	0.40	2.0 (0.4–3.5)	0.01	1.3 (0.3–2.2)	0.01
Growth deviation week 37, %	−5.1 (−6.6 to −3.6)	−2.7 (−3.3 to −2.0)	−0.4 (−2.1 to 1.4)	−2.4 (−4.0 to −0.8)	0.004	2.3 (0.4–4.2)	0.02	2.4 (1.2–3.5)	≤ 0.001
Change in growth deviation week 25–33, %	−2.1 (−3.3 to −1.0)	−1.5 (−2.0 to −1.0)	−0.8 (−2.0 to 0.4)	−0.6 (−1.9 to 0.6)	0.31	0.7 (−0.6 to 2.0)	0.28	0.7 (−0.2 to 1.5)	0.11
Change in growth deviation week 33–37, %	−0.5 (−1.6 to 0.6)	1.1 (0.7–1.5)	1.6 (0.4–2.8)	−1.6 (−2.8 to −0.4)	0.007	0.5 (−0.7 to 1.8)	0.42	1.1 (0.3–1.9)	0.008
Change in growth deviation week 25–37, %	−2.8 (−4.3 to −1.4)	−0.4 (−0.9–0.2)	0.8 (−0.9 to 2.5)	−2.5 (−4.0 to −0.9)	0.002	1.1 (−0.6 to 2.9)	0.21	1.8 (0.7–2.9)	0.001

Linear regression fitted by generalized linear models after multiple imputation. Covariates included: previous low birthweight child, maternal education, pre-pregnancy body mass index (transformed) and age (transformed), weight gain during pregnancy, smoking, gestational age (not for growth deviations), time between ultrasound dates (only for growth deviation intervals) and sex.
[a]Nonidentity link function.

Table 4. Birth outcome by Δ glucose group; mean/risk, comparison between groups and linear trend test (Low–Medium–High).

	Mean (95% CI)			Mean difference (95% CI), *p*-value				Mean (95% CI), *p*-value	
	Low (n = 103)	Medium (n = 661)	High (n = 91)	Low vs. Medium		High vs. Medium		Linear trend test	
Birthweight, g[a]	3517.6 (3451.6–3583.6)	3585.6 (3555.8–3615.4)	3655.8 (3578.1–3733.5)	−68.0 (−140.9 to 4.9)	0.07	70.2 (−13.0 to 153.3)	0.10	69.0 (18.2–119.9)	0.008
Birth length, cm	50.5 (50.2–50.8)	50.5 (50.4–50.7)	50.8 (50.4–51.1)	0.0 (−0.3 to 0.3)	0.94	0.2 (−0.1 to 0.6)	0.20	0.1 (−0.1 to 0.4)	0.29
Head circumference, cm	35.1 (34.9–35.3)	35.2 (35.1–35.3)	35.5 (35.3–35.7)	−0.1 (−0.3 to 0.1)	0.46	0.3 (0.1–0.5)	0.002	0.2 (0.1–0.3)	0.006
Triceps skin fold thickness, mm[a]	4.5 (4.3–4.6)	4.7 (4.7–4.8)	4.9 (4.7–5.1)	−0.3 (−0.4 to −0.1)	0.002	0.2 (−0.1 to 0.4)	0.16	0.2 (0.1–0.4)	0.002
Ponderal index, g*100/cm^3	2.7 (2.7–2.8)	2.8 (2.7–2.8)	2.8 (2.7–2.8)	−0.1 (−0.1 to 0.0)	0.02	0.0 (0.0–0.1)	0.71	0.0 (0.0–0.1)	0.03
Placental weight, g[a]	630.0 (607.3–652.6)	627.1 (618.0–636.1)	664.3 (641.9–686.7)	2.9 (−21.5 to 27.3)	0.82	37.2 (13.2–61.2)	0.002	16.2 (−0.1 to 32.6)	0.05
Gestational length, days	283.4 (281.8–284.9)	283.6 (283.0–284.2)	283.1 (281.4–284.8)	−0.2 (−1.9 to 1.4)	0.78	−0.5 (−2.3 to 1.3)	0.59	−0.1 (−1.3 to 1.0)	0.85
	Risk (95% CI)			Risk ratio (95% CI), *p*-value					
IUGR, %	18 (10–25)	11 (9–14)	10 (4, −17)	1.58 (0.98–2.54)	0.06	0.93 (0.50–1.74)	0.83	0.74 (0.49–1.09)	0.13
SGA, %	19 (12–26)	13 (10–15)	8 (3–12)	1.46 (0.96–2.24)	0.08	0.60 (0.31–1.15)	0.13	0.65 (0.47–0.90)	0.009

Linear (continuous) and Poisson (binary) regression fitted by generalized linear models after imputation. Covariates included: previous low birthweight child, maternal education, pre-pregnancy body mass index (transformed) and age (transformed), weight gain during pregnancy, smoking, gestational age (not for gestational length nor the frequencies) and sex.
IUGR, intrauterine growth restriction; SGA, small for gestational age.
[a]Nonidentity link function.

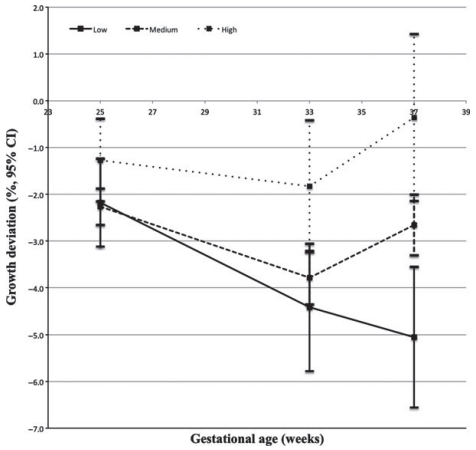


Figure 2. Mean deviation (%) from expected fetal weight (95% CI) by Δ glucose group.

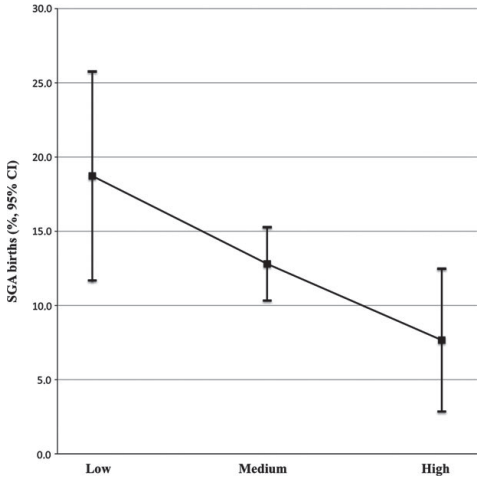


Figure 3. Prevalence (%) of small-for-gestational age (SGA) births (95% CI) by Δ glucose group.

they had in week 25 (−0.4% in M and 0.8% in H). Group L, however, deviated by an extra −2.8 percentage points in week 37 compared with week 25.

Birthweight and ponderal index followed the same order as predicted during pregnancy. Triceps skin-fold thicknesses also showed a strong positive correlation with Δ glucose. Head circumference was greater in group H than in the other two groups. Although similar in groups L and M, the placental weight was higher in group H.

Table 4 also shows the distribution of IUGR and SGA frequencies. There was a 58% increased risk for IUGR in group L compared with group M. The prevalence of SGA births followed a negative linear trend between the ordered Δ glucose groups (Figure 3).

The association between low Δ glucose and impaired fetal growth was less evident when we restricted the analysis to the 10% random sample (see Supporting information, Table S3). At birth there were only minor differences between groups L and M (see Supporting information, Table S4).

All covariates under study had a substantial impact on fetal growth, except for parity and gestational age at the time of the OGTT. Hence, these two covariates were excluded from the analyses. The univariate analyses and multivariate analyses before and after imputation were comparable (see Supporting information, Tables S1 and S2).

Discussion

The results support our hypothesis that there is a positive correlation between Δ glucose and fetal growth. From week 25 onward, group H had a higher fetal weight than the other two groups. Group L steadily declined from its expected fetal weight during the last trimester and was the only group with a substantial change in weight deviation from week 25 to week 37.

All groups had a negative expected value at all time points. This may be explained by the higher birthweight at the time of sampling of the reference population, and by the high-risk population in our study (12,19).

The hyperinsulinemic euglycemic glucose clamp technique is widely accepted as the reference standard for assessing insulin sensitivity (10). As this test is relatively complex, other techniques that indirectly measure insulin sensitivity have emerged. Still, most of them involve both levels of glucose and insulin in a fasting state and preferably also after a glucose challenge (10). An argument for the use of Δ glucose rather than insulin sensitivity is therefore that the former is more accessible and easier to manage.

It seems plausible that Δ glucose serves as a proxy for insulin resistance, and even more so for glucose tolerance; almost all (96%) women included in group H fulfilled the criteria for either GDM or DM. A high fasting glucose may contribute to a lower Δ glucose value. Nevertheless, the fasting glucose value was comparable between the groups. Still, we emphasize that Δ glucose is an approximation of glucose tolerance. The finding of heavier fetuses in the diabetogenic group H was in accordance with the accepted understanding that GDM and DM are associated with accelerated fetal growth (20).

In a small, high-risk population, Langer et al. found that hypoglycemic pregnancies had a 10-fold increased risk for SGA births compared with controls (21). In a similar comparison, Abell and Beischer studied 2000 low-risk pregnancies and observed a doubled risk for low birthweight in hypoglycemic pregnancies (22). Varying exposure and outcome definitions, as well as different populations under study, may explain why these reports showed a more pronounced association than we did. For instance, in the former, hypoglycemic pregnancies were compared with normal and hyperglycemic pregnancies combined, which may have exaggerated the effect estimate (21). Weissman et al. found no association between hypoglycemia and SGA births and concluded that postprandial hypoglycemia should be regarded as a normal phenomenon in pregnancy (23). A notable difference in that study, compared with Langer et al. (21), Abell et al. (22) and our study, is that the OGTT was performed in the second trimester and not the third. Failure to increase the glucose intolerance in the third trimester in relation to the fetal growth spurt may have more pronounced effects on fetal growth than an elevated glucose tolerance in the second trimester.

The insulin levels among women giving birth to growth-restricted babies have repeatedly been found to be lower than those among controls (24,25). Furthermore, in two small studies and a recent, large one, increased insulin sensitivity was found in growth restricted pregnancies (9,26,27). If our hypothesis that many of the women in group L were too insulin sensitive is correct, our findings are in accordance with these studies.

Two distinct strengths of our study are the repeated ultrasound measurements and the large study population. In our attempts to identify IUGR, a longitudinal design and extensive fetal weight data allowed us to estimate individual fetal growth patterns, which is much preferred over cross-sectional measures (28). The OGTT was performed regardless of risk factors for GDM, providing us with data at both ends of the insulin resistance range.

A limitation of the present study is that we lacked data on maternal insulin levels. We were therefore unable to correlate Δ glucose with insulin sensitivity. Furthermore, our study population differed from the general population in two important aspects. We had no nulliparous women, and our population was enriched with women with increased risk for giving birth to SGA babies. Analyses restricted to the 10% random sample revealed a weaker association between low Δ glucose and impaired growth. This could be due to a reduced sample size, smaller proportion of SGA births and a lower risk profile. Fetuses already at high risk for growth restriction may be particularly vulnerable to a compromised nutritional supply. Generalization of our findings should therefore be treated with caution. Data on placental function were unavailable to us and may have contributed to confounding because of its importance for both fetal growth and access to maternal blood glucose (29).

Given that they are replicated in future studies, our findings may have important implications for clinical practice. In today's antenatal care the sole reason for performing a routine OGTT is to diagnose and treat GDM or DM, thereby accepting values below these thresholds as normal (30). Our results indicate that if the 2-h glucose value from an OGTT in late pregnancy is roughly equal to or below the fasting value, they imply an increased risk for restricted growth, especially in a high-risk population. One may speculate that there may be therapeutic opportunities if excessive insulin sensitivity is found to cause IUGR (such as increase the insulin resistance). However, we underscore that such conclusions cannot be drawn from this study alone, and that more studies are needed before any clinical implementation. Besides challenging our findings in high-risk populations, we would suggest that observational studies of representative samples of the pregnant population should be performed to evaluate the predictive value of a low Δ glucose on IUGR. We would also encourage further studies of OGTT throughout all trimesters of pregnancy, both to evaluate the individual change in glucose tolerance and because an earlier identification of risk pregnancies may have clinical advantages. A systematic review on the current issue would be informative.

Conclusion

In summary, we found a positive relation between Δ glucose and fetal growth. Our results support the large body of evidence that pregnant women who are glucose intolerant often carry heavier babies. More important, we found an increased risk for impaired fetal growth in pregnancies with low Δ glucose, which has been sparsely reported previously. Our results advocate that more attention should be paid to pregnant women with an insufficient increase in glucose after a glucose challenge, with special emphasis on high-risk pregnancies.

Acknowledgements

We thank Marius Flasnes for helpful critical input; Hilde B. Lauritzen, MD, for assistance in estimating the fetal weight and weight deviations; Ivar Følling, MD, for including analyses of glucose and glycosylated hemoglobin as part of the original SGA study; and Lars Wibell, MD PhD, for sharing his knowledge on clinical aspects and pathophysiology.

Funding

The study was conducted as part of the commissioned activity of the publicly funded Norwegian University of Science and Technology.

References

1. Abel KM, Dalman C, Svensson AC, Susser E, Dal H, Idring S, et al. Deviance in fetal growth and risk of autism spectrum disorder. Am J Psychiatry. 2013;170:391–8.
2. Barker DJP, Clark PM. Fetal undernutrition and disease in later life. Rev Reprod. 1997;2:105–12.
3. Freinkel N. Banting Lecture 1980. Of pregnancy and progeny. Diabetes. 1980;29:1023–35.
4. Scholl TO, Sowers M, Chen X, Lenders C. Maternal glucose concentration influences fetal growth, gestation, and pregnancy complications. Am J Epidemiol. 2001;154:514–20.
5. Lockitch G. Clinical biochemistry of pregnancy. Crit Rev Clin Lab Sci. 1997;34:67–139.
6. Oliver MH, Harding JE, Breier BH, Evans PC, Gluckman PD. Glucose but not a mixed amino acid infusion regulates plasma insulin-like growth factor-I concentrations in fetal sheep. Pediatr Res. 1993;34:62–5.
7. Alberti KG, Zimmet PZ. Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. Diabet Med. 1998;15:539–53.
8. Persson B, Hanson U. Neonatal morbidities in gestational diabetes mellitus. Diabetes Care. 1998;21(Suppl 2):B79–84.
9. Catalano PM, Drago NM, Amini SB. Maternal carbohydrate metabolism and its relationship to fetal growth and body composition. Am J Obstet Gynecol. 1995;172:1464–70.
10. Muniyappa R, Lee S, Chen H, Quon MJ. Current approaches for assessing insulin sensitivity and resistance in vivo: advantages, limitations, and appropriate usage. Am J Physiol Endocrinol Metab. 2008;294:E15–26.
11. Bakketeig LS, Jacobsen G, Hoffman HJ, Lindmark G, Bergsjø P, Molne K, et al. Pre-pregnancy risk factors of small-for-gestational age births among parous women in Scandinavia. Acta Obstet Gynecol Scand. 1993;72:273–9.
12. Gjessing HK, Grøttum P, Eik-Nes SH. A direct method for ultrasound prediction of day of delivery: a new, population-based approach. Ultrasound Obstet Gynecol. 2007;30:19–27.
13. Gjessing H, Grøttum P, Økland I, Dreier J, Eik-Nes S. [eSnurra Growth – Description of Methods] [Internet]. 2012. pp. 1–8. Available online at: http://www.nsfm.no/esnurra (accessed October 4, 2013).
14. Bjerkedal T, Skjaerven R. Percentiles of birth weight and crown-heel length in relation to gestation period for single live births. Tidsskr Nor Laegeforen. 1980;100:1088–91.
15. Jørgensen LGM, Schytte T, Brandslund I, Stahl M, Petersen PH, Andersen B. Fasting and post-glucose load – reference limits for peripheral venous plasma glucose concentration in pregnant women. Clin Chem Lab Med. 2003;41:187–99.
16. Hebel JR, Fox NL, Sexton M. Dose–response of birth weight to various measures of maternal smoking during pregnancy. J Clin Epidemiol. 1988;41:483–9.
17. Zou G. A modified Poisson regression approach to prospective studies with binary data. Am J Epidemiol. 2004;159:702–6.
18. Helsedirektoratet. Nasjonale faglige retningslinjer. Diabetes. Forebygging, diagnostikk og behandling. [National guidance. Diabetes. Prevention, diagnosis and treatment]. (In Norwegian.) Oslo: Helsedirektoratet, 2009.
19. Norwegian Institute of Public Health. [Register of births in Norway in 2011] [Internet]. 2013. Available online ate: http://www.fhi.no/helseregistre/medisinsk-fodselsregister/ (accessed October 4, 2013).
20. Bernstein I. Fetal body composition. Curr Opin Clin Nutr Metab Care. 2005;8:613–7.
21. Langer O, Damus K, Maiman M, Divon M, Levy J, Bauman W. A link between relative hypoglycemia–hypoinsulinemia during oral glucose tolerance tests and intrauterine growth retardation. Am J Obstet Gynecol. 1986;155:711–6.
22. Abell D, Beischer N. The relationship between maternal glucose tolerance and fetal size at birth. Aust N Z J Obstet Gynaecol. 1976;16:1–4.
23. Weissman A, Solt I, Zloczower M, Jakobi P. Hypoglycemia during the 100-g oral glucose tolerance test: incidence and perinatal significance. Obstet Gynecol. 2005;105:1424–8.
24. Sokol RJ, Kazzi GM, Kalhan SC, Pillay SK. Identifying the pregnancy at risk for intrauterine growth retardation: possible usefulness of the intravenous glucose tolerance test. Am J Obstet Gynecol. 1982;143:220–3.
25. Khouzami VA, Ginsburg DS, Daikoku NH, Johnson JW. The glucose tolerance test as a means of identifying intrauterine growth retardation. Am J Obstet Gynecol. 1981;139:423–6.
26. Caruso A, Paradisi G, Ferrazzani S, Lucchese A, Moretti S, Fulghesu AM. Effect of maternal carbohydrate metabolism on fetal growth. Obstet Gynecol. 1998;92:8–12.
27. Dalfrà MG, Pacini G, Parretti E, Ragazzi E, Mello G, Lapolla A. Elevated insulin sensitivity and β-cell function during pregnancy in mothers of growth-restricted newborns. Am J Physiol Endocrinol Metab. 2011;301:E25–30.
28. Nasrat H. Use of ultrasound longitudinal data in the diagnosis of abnormal fetal growth. J Matern Fetal Med. 1997;6:209–14.
29. Baumann MU, Deborde S, Illsley NP. Placental glucose transfer and fetal growth. Endocrine. 2002;19:13–22.
30. Bartoli E, Fra GP. Carnevale Schianca GP. The oral glucose tolerance test (OGTT) revisited. Eur J Intern Med. 2011;22:8–12.

Supporting information

Additional Supporting Information may be found in the online version of this article:

Table S1. Fetal growth by Δ glucose group; mean comparison between groups and linear trend test. Univariate and complete case analysis.

Table S2. Birth outcome by Δ glucose group; mean/risk, comparison between groups and linear trend test. Univariate and complete case analysis.

Table S3. Fetal growth by Δ glucose group; mean, comparison between groups and linear trend test (Low–Medium–High). Random sample only.

Table S4. Birth outcome by Δ glucose group; mean/risk, comparison between groups and linear trend test (Low–Medium–High). Random sample only.

Table S1. Fetal growth by Δ glucose group; mean comparison between groups and linear trend test (Low-Medium-High). Univariate and complete case analysis.

	Low (n = 103)	Medium (n = 661)	High (n = 91)	Low vs. Medium		High vs Medium		Linear trend test		Missing
Univariate analysis	Mean (95% CI)			Mean Difference (95% CI), p-value				Mean (95% CI), p-value		
Estimated weight week 25, g†	780.6 (765.5, 795.8)	781.6 (774.5, 788.8)	790.1 (771.6, 808.5)	-1.0 (-17.7, 15.8)	p = 0.91	8.5 (-11.3, 28.3)	p = 0.40	4.6 (-7.4, 16.6)	p = 0.45	33
Estimated weight week 33, g	2 042.4 (1 999.4, 2 085.3)	2 033.3 (2 015.4, 2 051.2)	2 050.8 (2 008.1, 2 093.5)	9.1 (-37.4, 55.6)	p = 0.70	17.5 (-28.8, 63.8)	p = 0.46	3.8 (-26.6, 34.1)	p = 0.81	11
Estimated weight week 37, g	2 920.2 (2 857.2, 2 983.2)	2 995.2 (2 970.6, 3 019.7)	3 030.8 (2 963.0, 3 098.5)	-75.0 (-142.6, -7.4)	p = 0.03	35.6 (-36.5, 107.7)	p = 0.33	56.1 (9.9, 102.3)	p = 0.02	14
Growth deviation week 25, %	-2.0 (-3.1, -1.0)	-2.3 (-2.7, -1.9)	-1.2 (-2.1, -0.3)	0.3 (-0.8, 1.4)	p = 0.59	1.1 (0.1, 2.2)	p = 0.03	0.4 (-0.3, 1.1)	p = 0.28	19
Growth deviation week 33, %	-4.2 (-5.9, -2.6)	-3.8 (-4.4, -3.1)	-2.3 (-3.8, -0.8)	-0.5 (-2.2, 1.3)	p = 0.60	1.4 (-0.2, 3.1)	p = 0.09	0.9 (-0.2, 2.1)	p = 0.10	11
Growth deviation week 37, %	-4.8 (-6.6, -3.0)	-2.7 (-3.4, -1.9)	-0.8 (-2.7, 1.1)	-2.1 (-4.1, -0.2)	p = 0.03	1.9 (-0.2, 3.9)	p = 0.07	2.0 (0.7, 3.3)	p = 0.003	14
Change in growth deviation week 25-33, %	-2.2 (-3.3, -1.0)	-1.5 (-2.0, -1.0)	-1.2 (-2.4, 0.1)	-0.6 (-1.9, 0.7)	p = 0.34	0.4 (-1.0, 1.7)	p = 0.60	0.5 (-0.4, 1.4)	p = 0.26	25
Change in growth deviation week 33-37, %	-0.4 (-1.5, 0.7)	1.0 (0.6, 1.5)	1.6 (0.4, 2.8)	-1.4 (-2.6, -0.2)	p = 0.02	0.6 (-0.7, 1.8)	p = 0.36	1.0 (0.2, 1.8)	p = 0.01	20
Change in growth deviation week 25-37, %	-2.7 (-4.2, -1.3)	-0.5 (-1.1, 0.2)	0.3 (-1.4, 2.0)	-2.3 (-3.8, -0.7)	p = 0.005	0.7 (-1.1, 2.5)	p = 0.44	1.5 (0.4, 2.6)	p = 0.007	27
Complete case analysis	Mean (95% CI)			Mean Difference (95% CI), p-value				Mean (95% CI), p-value		
Estimated weight week 25, g†	784.7 (775.2, 794.3)	781.9 (777.6, 786.2)	792.8 (783.8, 801.7)	2.9 (-7.5, 13.2)	p = 0.59	10.9 (0.8, 21.0)	p = 0.04	4.0 (-2.8, 10.7)	p = 0.25	85
Estimated weight week 33, g	2 017.9 (1 988.5, 2 047.4)	2 035.0 (2 022.4, 2 047.7)	2 072.6 (2 042.5, 2 102.6)	-17.1 (-49.3, 15.1)	p = 0.30	37.5 (5.0, 70.1)	p = 0.02	27.0 (6.1, 47.9)	p = 0.01	52
Estimated weight week 37, g	2 918.8 (2 870.6, 2 966.9)	3 001.9 (2 981.1, 3 022.6)	3 061.5 (3 005.4, 3 117.6)	-83.1 (-135.6, -30.6)	p = 0.002	59.6 (-0.2, 119.4)	p = 0.05	71.8 (35.1, 108.5)	p = <0.001	63
Growth deviation week 25, %	-2.2 (-3.2, -1.2)	-2.2 (-2.6, -1.8)	-1.1 (-2.0, -0.2)	0.0 (-1.0, 1.1)	p = 0.94	1.1 (0.1, 2.1)	p = 0.03	0.5 (-0.2, 1.2)	p = 0.13	71
Growth deviation week 33, %	-4.5 (-5.9, -3.1)	-3.7 (-4.2, -3.1)	-1.8 (-3.2, -0.4)	-0.8 (-2.3, 0.7)	p = 0.29	1.9 (0.3, 3.4)	p = 0.02	1.3 (0.3, 2.3)	p = 0.009	52
Growth deviation week 37, %	-5.2 (-6.7, -3.6)	-2.4 (-3.1, -1.8)	-0.5 (-2.3, 1.3)	-2.7 (-4.4, -1.0)	p = 0.002	1.9 (0.0, 3.9)	p = 0.05	2.3 (1.1, 3.5)	p = <0.001	63
Change in growth deviation week 25-33, %	-2.3 (-3.5, -1.1)	-1.5 (-2.0, -1.0)	-0.9 (-2.1, 0.3)	-0.8 (-2.1, 0.5)	p = 0.22	0.6 (-0.7, 1.9)	p = 0.37	0.7 (-0.1, 1.6)	p = 0.10	64
Change in growth deviation week 33-37, %	-0.4 (-1.5, 0.8)	1.0 (0.6, 1.5)	1.5 (0.3, 2.7)	-1.4 (-2.6, -0.2)	p = 0.02	0.5 (-0.8, 1.8)	p = 0.47	1.0 (0.1, 1.8)	p = 0.02	67
Change in growth deviation week 25-37, %	-3.0 (-4.4, -1.5)	-0.4 (-0.9, 0.2)	0.4 (-1.3, 2.1)	-2.6 (-4.2, -1.0)	p = 0.001	0.8 (-1.0, 2.5)	p = 0.41	1.7 (0.6, 2.8)	p = 0.003	74

Linear regression fitted by generalized linear models. Covariates included in the complete case analyses: Previous low birth weight child, maternal education, pre-pregnancy BMI (transformed) and age (transformed), weight gain during pregnancy, smoking, gestational age (not for growth deviations), time between ultrasound dates (only for growth deviation intervals) and sex.

† Non-identity link function. CI - confidence interval.

Table S2. Birth outcome by Δ glucose group; mean/risk, comparison between groups and linear trend test (Low-Medium-High). Univariate and complete case analysis.

	Low (n = 103)	Medium (n = 661)	High (n = 91)	Low vs Medium		High vs Medium		Linear trend test		Missing
Univariate analysis	Mean (95% CI)			Mean Difference (95%CI), p-value				Mean (95% CI), p-value		
Birth weight, g†	3 523.3 (3 436.7, 3 609.9)	3 588.5 (3 552.1, 3 625.0)	3 627.2 (3 533.8, 3 720.6)	-65.2 (-159.1, 28.7)	p = 0.17	38.7 (-61.6, 138.9)	p = 0.45	52.5 (-11.0, 116.0)	p = 0.11	0
Birth length, cm	50.5 (50.2, 50.9)	50.6 (50.4, 50.7)	50.7 (50.3, 51.2)	0.0 (-0.4, 0.4)	p = 0.97	0.2 (-0.3, 0.6)	p = 0.52	0.1 (-0.2, 0.4)	p = 0.58	8
Head circumference, cm	35.1 (34.9, 35.4)	35.2 (35.1, 35.3)	35.4 (35.2, 35.6)	-0.1 (-0.3, 0.2)	p = 0.59	0.2 (0.0, 0.5)	p = 0.09	0.1 (0.0, 0.3)	p = 0.10	7
Triceps skin fold thickness, mm†	4.5 (4.3, 4.6)	4.7 (4.7, 4.8)	4.9 (4.7, 5.1)	-0.2 (-0.4, -0.1)	p = 0.007	0.2 (0.0, 0.4)	p = 0.10	0.2 (0.1, 0.4)	p = 0.001	18
Ponderal index, g*100/cm^3	2.7 (2.7, 2.8)	2.8 (2.7, 2.8)	2.8 (2.7, 2.8)	-0.1 (-0.1, 0.0)	p = 0.06	0.0 (-0.1, 0.1)	p = 0.91	0.0 (0.0, 0.1)	p = 0.11	8
Placental weight, g†	634.2 (609.0, 659.3)	627.2 (617.5, 636.8)	657.1 (634.3, 679.9)	7.0 (-20.0, 33.9)	p = 0.61	29.9 (5.2, 54.7)	p = 0.02	10.8 (-6.6, 28.2)	p = 0.23	59
Gestational length, days	283.6 (282.0, 285.1)	283.6 (282.9, 284.2)	283.3 (281.7, 284.9)	0.0 (-1.7, 1.7)	p = 0.99	-0.2 (-2.0, 1.5)	p = 0.79	-0.1 (-1.2, 1.0)	p = 0.83	5
	Risk (95% CI)			Risk Ratio (95% CI), p-value				Risk Ratio (95% CI), p-value		
IUGR, %	18 (10, 26)	11 (9, 14)	11 (5, 18)	1.62 (1.01, 2.60)	p = 0.05	1.01 (0.54, 1.89)	p = 0.97	0.75 (0.49, 1.13)	p = 0.16	27
SGA, %	18 (11, 26)	13 (10, 15)	9 (3, 15)	1.47 (0.93, 2.31)	p = 0.10	0.70 (0.35, 1.40)	p = 0.31	0.69 (0.48, 0.98)	p = 0.04	0
Complete case analysis	Mean (95% CI)			Mean Difference (95% CI), p-value				Mean (95% CI), p-value		
Birth weight, g†	3 541.8 (3 473.0, 3 610.7)	3 615.2 (3 582.9, 3 647.5)	3 675.6 (3 594.4, 3 756.9)	-73.3 (-149.8, 3.1)	p = 0.06	60.5 (-26.9, 147.8)	p = 0.18	67.2 (14.0, 120.4)	p = 0.01	100
Birth length, cm	50.6 (50.3, 50.9)	50.7 (50.5, 50.8)	50.9 (50.6, 51.3)	0.0 (-0.4, 0.3)	p = 0.90	0.3 (-0.1, 0.6)	p = 0.17	0.1 (-0.1, 0.4)	p = 0.24	106
Head circumference, cm	35.2 (35.0, 35.4)	35.3 (35.2, 35.3)	35.6 (35.4, 35.8)	-0.1 (-0.3, 0.1)	p = 0.37	0.3 (0.1, 0.5)	p = 0.003	0.2 (0.1, 0.3)	p = 0.004	105
Triceps skin fold thickness, mm†	4.5 (4.3, 4.6)	4.7 (4.7, 4.8)	4.9 (4.7, 5.2)	-0.3 (-0.4, -0.1)	p = 0.003	0.2 (-0.1, 0.4)	p = 0.12	0.2 (0.1, 0.4)	p = 0.001	113
Ponderal index, g*100/cm^3	2.7 (2.7, 2.8)	2.8 (2.8, 2.8)	2.8 (2.7, 2.8)	-0.1 (-0.1, 0.0)	p = 0.03	0.0 (-0.1, 0.1)	p = 0.93	0.0 (0.0, 0.1)	p = 0.08	106
Placental weight, g†	634.5 (611.0, 658.1)	632.5 (622.7, 642.3)	657.9 (636.6, 679.3)	2.1 (-23.3, 27.5)	p = 0.87	25.5 (1.9, 49.0)	p = 0.03	11.1 (-5.1, 27.3)	p = 0.18	146
Gestational length, days	283.8 (282.3, 285.4)	284.3 (283.7, 284.9)	283.4 (281.7, 285.1)	-0.5 (-2.1, 1.2)	p = 0.58	-0.9 (-2.7, 0.9)	p = 0.33	-0.2 (-1.3, 1.0)	p = 0.75	100
	Risk (95% CI)			Risk Ratio (95% CI), p-value				Risk Ratio (95% CI), p-value		
IUGR, %	18 (10, 27)	11 (9, 14)	11 (4, 17)	1.64 (1.01, 2.67)	p = 0.05	0.94 (0.50, 1.75)	p = 0.84	0.72 (0.48, 1.09)	p = 0.12	74
SGA, %	19 (12, 27)	13 (10, 15)	7 (2, 12)	1.53 (0.97, 2.40)	p = 0.07	0.55 (0.26, 1.17)	p = 0.12	0.62 (0.44, 0.88)	p = 0.007	96

Linear (continuous) and Poisson (binary) regression fitted by generalized linear models.Covariates included in the complete case analyses: Previous low birth weight child, maternal education, pre-pregnancy BMI (transformed) and age (transformed), weight gain during pregnancy, smoking, gestational age (not for gestational length nor frequencies) and sex.

† Non-identity link function. CI - confidence interval; IUGR - intrauterine growth restriction; SGA - small for gestational age.

Table S3. Fetal growth by Δ glucose group; mean, comparison between groups and linear trend test (Low-Medium-High). Random sample only.

	Low (n = 25)	**Medium (n = 209)**	**High (n = 23)**	**Low vs Medium**		**High vs Medium**		**Linear trend test**	
	Mean (95% CI)			Mean Difference (95% CI), p-value				Mean (95% CI), p-value	
Estimated weight week 25, g†	785.9 (770.4, 801.4)	773.8 (768.3, 779.3)	793.6 (774.9, 812.3)	12.1 (-4.1, 28.4)	p = 0.14	19.8 (0.3, 39.3)	p = 0.05	4.7 (-8.8, 18.2)	p = 0.49
Estimated weight week 33, g	2 052.4 (2 005.3, 2 099.5)	2 049.8 (2 030.1, 2 069.4)	2 090.3 (2 031.4, 2 149.2)	2.6 (-48.4, 53.7)	p = 0.92	40.5 (-21.3, 102.3)	p = 0.20	18.0 (-19.8, 55.8)	p = 0.35
Estimated weight week 37, g	2 968.9 (2 899.7, 3 038.1)	3 044.9 (3 011.8, 3 078.0)	3 062.9 (2 956.7, 3 169.2)	-75.9 (-152.9, 1.1)	p = 0.05	18.1 (-94.2, 130.3)	p = 0.75	48.0 (-14.4, 110.3)	p = 0.13
Growth deviation week 25, %	-0.5 (-2.2, 1.2)	-1.8 (-2.4, -1.2)	0.6 (-1.5, 2.8)	1.3 (-0.5, 3.1)	p = 0.16	2.4 (0.2, 4.7)	p = 0.04	0.5 (-0.9, 1.9)	p = 0.50
Growth deviation week 33, %	-2.4 (-4.7, -0.2)	-2.6 (-3.5, -1.7)	-0.6 (-3.4, 2.1)	0.2 (-2.2, 2.6)	p = 0.89	2.0 (-1.0, 4.9)	p = 0.19	0.9 (-0.9, 2.6)	p = 0.35
Growth deviation week 37, %	-3.6 (-5.9, -1.4)	-0.9 (-2.0, 0.2)	-0.5 (-4.0, 2.9)	-2.7 (-5.2, -0.2)	p = 0.03	0.4 (-3.3, 4.0)	p = 0.84	1.6 (-0.4, 3.6)	p = 0.12
Change in growth deviation week 25-33, %	-1.7 (-4.2, 0.8)	-0.8 (-1.6, 0.0)	-1.6 (-3.4, 0.1)	-0.9 (-3.6, 1.7)	p = 0.48	-0.9 (-2.8, 1.0)	p = 0.38	0.1 (-1.5, 1.7)	p = 0.92
Change in growth deviation week 33-37, %	-1.0 (-3.4, 1.3)	1.6 (0.9, 2.4)	0.3 (-1.8, 2.4)	-2.7 (-5.1, -0.2)	p = 0.04	-1.4 (-3.6, 0.9)	p = 0.23	0.7 (-0.9, 2.3)	p = 0.40
Change in growth deviation week 25-37, %	-3.1 (-5.9, -0.3)	0.9 (-0.1, 1.8)	-1.2 (-4.0, 1.5)	-3.9 (-6.9, -1.0)	p = 0.009	-2.1 (-5.0, 0.9)	p = 0.17	1.0 (-1.0, 3.1)	p = 0.32

Linear regression fitted by generalized linear models after multiple imputation. Covariates included: Previous low birth weight child, maternal education, pre-pregnancy BMI (transformed) and age (transformed), weight gain during pregnancy, smoking, gestational age (not for growth deviations), time between ultrasound dates (only for growth deviation intervals) and sex.

† Non-identity link function. CI - confidence interval.

Table S4. Birth outcome by Δ glucose group; mean/risk, comparison between groups and linear trend test (Low-Medium-High). Random sample only.

	Low (n = 25)	**Medium (n = 209)**	**High (n = 23)**	**Low vs Medium**		**High vs Medium**		**Linear trend test**	
	Mean (95% CI)			Mean Difference (95%CI), p-value				Mean (95% CI), p-value	
Birth weight, g†	3 657.2 (3 524.5, 3 789.9)	3 679.1 (3 625.1, 3 733.1)	3 701.6 (3 554.5, 3 848.7)	-21.9 (-165.0, 121.2)	p = 0.76	22.5 (-136.4, 181.5)	p = 0.78	22.2 (-77.3, 121.7)	p = 0.66
Birth length, cm	51.3 (50.7, 51.9)	50.9 (50.7, 51.1)	51.7 (50.8, 52.5)	0.4 (-0.2, 1.1)	p = 0.19	0.8 (-0.1, 1.7)	p = 0.08	0.2 (-0.4, 0.7)	p = 0.56
Head circumference, cm	35.5 (35.1, 35.9)	35.4 (35.2, 35.5)	35.7 (35.4, 36.1)	0.1 (-0.3, 0.5)	p = 0.60	0.4 (0.0, 0.7)	p = 0.03	0.1 (-0.1, 0.4)	p = 0.35
Triceps skin fold thickness, mm†	4.5 (4.2, 4.7)	4.8 (4.6, 4.9)	4.9 (4.4, 5.3)	-0.3 (-0.6, 0.0)	p = 0.05	0.1 (-0.4, 0.5)	p = 0.75	0.2 (-0.1, 0.5)	p = 0.14
Ponderal index, g*100/cm^3	2.7 (2.6, 2.8)	2.8 (2.8, 2.8)	2.7 (2.6, 2.8)	-0.1 (-0.2, 0.0)	p = 0.01	-0.1 (-0.2, 0.0)	p = 0.20	0.0 (-0.1, 0.1)	p = 0.93
Placental weight, g†	650.7 (607.5, 693.8)	639.1 (623.2, 655.0)	646.9 (613.3, 680.4)	11.5 (-34.2, 57.3)	p = 0.62	7.7 (-29.5, 45.0)	p = 0.68	-2.4 (-30.8, 25.9)	p = 0.87
Gestational length, days	283.8 (280.9, 286.7)	284.0 (283.0, 284.9)	284.4 (281.1, 287.7)	-0.2 (-3.2, 2.8)	p = 0.90	0.4 (-3.0, 3.8)	p = 0.82	0.3 (-1.9, 2.5)	p = 0.79
	Risk (95% CI)			Risk Ratio (95% CI), p-value				Risk Ratio (95% CI), p-value	
IUGR, %	24 (3, 44)	10 (6, 14)	11 (1, 21)	2.39 (0.89, 6.39)	p = 0.08	1.15 (0.41, 3.24)	p = 0.80	0.66 (0.29, 1.49)	p = 0.32
SGA, %	9 (0, 19)	10 (6, 15)	4 (-1, 9)	0.90 (0.29, 2.81)	p = 0.86	0.40 (0.10, 1.63)	p = 0.20	0.70 (0.40, 1.21)	p = 0.20

Linear (continuous) and Poisson (binary) regression fitted by generalized linear models after imputation. Covariates included: Previous low birth weight child, maternal education, pre-pregnancy BMI (transformed) and age (transformed), weight gain during pregnancy, smoking, gestational age (not for gestational length nor the frequencies) and sex.

† Non-identity link function. CI - confidence interval; IUGR - intrauterine growth restriction; SGA - small for gestational age.

Paper II

Is not included due to copyright

Paper III

Maternal vitamin B12 in pregnancy and risk of preterm birth and low birth weight: A systematic review and individual participant data meta-analysis

Tormod Rogne,[1] Myrte J Tielemans,[2,3] Mary Foong-Fong Chong,[4,5,6] Chittaranjan S Yajnik,[7] Ghattu V Krishnaveni,[8] Lucilla Poston,[9] Vincent W V Jaddoe,[3,10] Eric A P Steegers,[11] Suyog Joshi,[7] Yap-Seng Chong,[4,5,12] Keith M Godfrey,[13] Fabian Yap,[14,15,16] Raquel Yahyaoui,[17] Tinku Thomas,[18] Gry Hay,[19] Marije Hogeveen,[20] Ahmet Demir,[21] Ponnusamy Saravanan,[22] Eva Skovlund,[1,23] Marit P Martinussen,[24,25] Geir W Jacobsen,[1] Oscar H Franco,[2] Michael B Bracken,[26] Kari R Risnes.[1,27]*

[1] *Affiliation*: Department of Public Health and General Practice, Norwegian University of Science and Technology, Trondheim, Norway; *Address*: NTNU, Det medisinske fakultet, Institutt for samfunnsmedisin, postboks 8905, 7491 Trondheim, Norway.

[2] *Affiliation*: Department of Epidemiology, Erasmus MC, Rotterdam, the Netherlands; *Address*: P.O. Box 2040, Rotterdam CA 3000 CA Rotterdam, the Netherlands.

[3] *Affiliation*: The Generation R Study Group, Erasmus MC, Rotterdam, the Netherlands; *Address*: P.O. Box 2040, Rotterdam CA 3000 CA Rotterdam, the Netherlands.

[4] *Affiliation*: Singapore Institute for Clinical Sciences, A*STAR, Singapore; *Address*: Singapore Institute for Clinical Sciences, Brenner Centre for Molecular Medicine, 30 Medical Drive, Singapore 117609.

[5] *Affiliation*: Clinical Nutrition Research Centre, Singapore Institute for Clinical Sciences, A*STAR, Singapore.

[6] *Affiliation*: Department of Paediatrics,Yong Loo Lin School of Medicine, National University of Singapore, Singapore.

[7] *Affiliation*: Kamalnayan Bajaj Diabetology Research Centre, King Edward Memorial Hospital Research Centre, Pune, India; *Address*: Kamalnayan Bajaj Diabetology Research Centre,6th floor Banoo Coyaji Building, King Edward Memorial Hospital Research Centre, Rasta Peth, Pune-411011, Maharashtra state, India.

[8] *Affiliation*: Epidemiology Research Unit, CSI Holdsworth Memorial Hospital, Mysore, India; *Address*: Epidemiology Research Unit, CSI Holdsworth Memorial Hospital, Mandi Mohalla, Mysore, Karnataka, India.

[9] *Affiliation*: Division of Women's Health, King's College London and King's Health Partners, London, United Kingdom; *Address*: Division of Women's Health, Women's Health Academic Centre, King's College London and King's Health Partners, 10th floor, North Wing, St.Thomas' Hospital, London SE1 7EH, United Kingdom.

[10] *Affiliation*: Department of Pediatrics, Erasmus MC, Rotterdam, the Netherlands; *Address*: P.O. Box 2040, Rotterdam CA 3000 CA Rotterdam, the Netherlands.

[11] *Affiliation*: Department of Obstetrics and Gynecology, Erasmus MC, Rotterdam, the Netherlands; *Address*: P.O. Box 2040, Rotterdam CA 3000 CA Rotterdam, the Netherlands.

[12] *Affiliation*: Department of Obstetrics and Gynaecology, Yong Loo Lin School of Medicine, National University of Singapore and National University Health System, Singapore.

[13] *Affiliation*: Medical Research Council Lifecourse Epidemiology Unit and NIHR Southampton Biomedical Research Centre, University of Southampton and University Hospital Southampton NHS Foundation Trust, Southampton, United Kingdom; *Address*: MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton General Hospital, Mailpoint 95, Southampton SO16 6YD, United Kingdom.

[14] *Affiliation*: Department of Paediatrics, KK Women's and Children's Hospital, Singapore; *Address*: Department of Paediatrics, KK Women's and Children's Hospital 100, Bukit Timah Road, Singapore 229899.

[15] *Affiliation*: Duke-NUS Medical School, Singapore.

[16] *Affiliation*: Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore.

[17] *Affiliation*: Clinical Laboratory and Newborn Screening Center of Oriental Andalucía, Málaga Regional Hospital and Instituto de Investigación Biomédica de Málaga, IBIMA, Málaga, Spain; *Address*: Laboratorio de Metabolopatías. Hospital Materno Infantil, Av. Arroyo de los Angeles s/n 29011 Málaga, Spain.

[18] *Affiliation*: Division of Biostatistics and Epidemiology, St John's Research Institute, Bangalore, India; *Address*: 100 Feet Rd, John Nagar, Koramangala, Bengaluru, Karnataka 560034, India.

[19] *Affiliation*: The Norwegian Directorate of Health, Oslo, Norway; *Address*: Helsedirektoratet, Postboks 7000, St. Olavs plass, 0130 Oslo, Norway.

[20] *Affiliation*: Department of Pediatrics, Radboud University Medical Center, Nijmegen, the Netherlands; *Address*: Radboudumc Amalia kinderziekenhuis, Postbus 9101, 6500 HB Nijmegen (804), Geert Grooteplein 10 (route 804), 024-3614430, the Netherlands.

[21] *Affiliation*: Department of Obstetrics and Gynecology, Elazig Training and Research Hospital, Elazig, Turkey; *Address*: Elazig Egitim ve Arastirma Hastanesi, Kadın Hastalıkları ve Dogum Klinigi, Merkez/Elazig/Turkey.

[22] *Affiliation*: Warwick Medical School, University of Warwick, United Kingdom; George Eliot Hospital, Nuneaton, United Kingdom; *Address*: Gibett Hill, Warwick Medical School, University of Warwick, CV4 7AL, United Kingdom.

[23] *Affiliation*: Norwegian Institute of Public Health, Oslo, Norway.

[24] *Affiliation*: Department of Obstetrics and Gynecology, St. Olav's University Hospital, Trondheim, Norway.

[25] *Affiliation*: Department of Laboratory Medicine, Children's and Women's Health, Norwegian University of Science and Technology, Trondheim, Norway; *Address*: NTNU, Faculty of Medicine, Department for laboratory Medicine, Children's and Women's Health, P.O. Box 8905, Medisinsk teknisk forskningssenter, 7491 Trondheim.

[26] *Affiliation*: Yale Center for Perinatal, Pediatric and Environmental Epidemiology, Yale University, CT, USA; *Address*: Yale School of Public Health, PO Box 208034, 60 College Street, New Haven, CT, 06520-8034, USA.

[27] *Affiliation*: Department of Pediatrics, St. Olav's University Hospital, Trondheim, Norway

* Corresponding author. Email: tormod.rogne@ntnu.no and trogne@gmail.com; telephone number: (+47) 971 90 271; fax number: (+47) 73 59 75 77.

Funding source: This work was supported by the Norwegian University of Science and Technology as part of the Ph.D. in Medicine-programme. The funding source was not involved in the design of the study, statistical analysis and results interpretation. The researchers were independent from the funders.

Word count: Manuscript: 3,894, Abstract: 195

Number of Figures: 3 (+3 as Web Material). **Number of Tables:** 2 (+6 as Web Material).

Keywords: Low birth weight, preterm birth, pregnancy, systematic review, vitamin B12

Competing interest statement:

Keith M Godfrey and Yap-Seng Chong have received reimbursement for speaking at conferences sponsored by companies selling nutritional products. Keith M Godfrey and Yap-Seng Chong are part of an academic consortium that has received funding from Abbott Nutrition, Nestec and Danone. Oscar H Franco has received funding from Nestlé for research purposes. The other authors declare no support from any organisation for the submitted work, no financial relationships with any organisations that might have an interest in the submitted work in the previous three years, no other relationships or activities that could appear to have influences the submitted work.

Data sharing statement

All participating studies agreed on the use of data material for the purpose of this systematic review.

Ethics committee approval

This study was approved by the Regional Committee for Medical and Health Research Ethics, Norway. The studies included in this review were approved by their respective regional ethics committees.

Transparency declaration

The lead author (TR) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported, that no important aspects of the study have been omitted, and that any discrepancies from the study as planned have been explained.

ABSTRACT

Vitamin B12-deficiency in pregnancy is prevalent, and has been associated with lower birth weight (birth weight <2,500 g) and preterm birth (length of gestation <37 weeks). Nevertheless, current evidence is contradictory. We performed a systematic review and an individual participant data meta-analysis to evaluate the associations between maternal serum or plasma vitamin B12 (B12) concentration in pregnancy and offspring birth weight and length of gestation. Twenty-two eligible studies were identified (11,993 observations). Eighteen studies were included in the meta-analysis (11,216 observations). No linear association was observed between maternal B12 levels in pregnancy and birth weight, but B12-deficiency (<148 pmol/L) was associated with an increased risk of newborn low birth weight (adjusted risk ratio (RR) 1.15 (95% confidence interval (CI) 1.01, 1.31)). There was a linear association between maternal levels of B12 and preterm birth (adjusted RR for preterm birth was 0.89 (95% CI 0.82, 0.97) per one standard deviation increase in B12). Accordingly, B12-deficiency was associated with increased risk of preterm birth (adjusted RR 1.21 (95% CI 0.99, 1.49). Lower maternal B12 in pregnancy increased the risk of preterm birth. This finding supports the conduct of randomized controlled trials of vitamin B12 supplementation in pregnancy.

MANUSCRIPT

Globally, preterm birth and low birth weight (LBW) cause over a third of the 2.9 million neonatal deaths each year, and prevention of these events is important to reduce under-five year mortality (1,2). The etiology of preterm birth, however, is complex, and few interventions have been successful in preventing it (3).

Vitamin B12 (B12) is a vitamin with metabolic roles closely related to folate and homocysteine, and is found in animal-derived foods only (4). It is important for the synthesis (5) and methylation (6) of DNA, and plays a role in the energy production of the cell (7). It has been hypothesized that B12 may affect placentation and fetal growth (8). B12-deficiency may affect over three quarters of some pregnant populations (9).

Few supplementation-studies of B12 in pregnancy have been undertaken to assess possible effects on birth weight and length of gestation. However, a recent meta-analysis concluded that multiple-micronutrient supplementation may reduce the risk of LBW and the number of stillbirths, but not preterm birth or neonatal mortality (10). Thus, a more targeted micronutrient supplementation practice may be warranted.

The aim of this systematic review and individual participant data (IPD) meta-analysis was to study whether maternal serum or plasma B12 levels in pregnancy may be associated with birth weight and length of gestation. Individual studies have reported conflicting results. A recent systematic review that included traditional meta-analyses was unable to conclude whether maternal B12 levels were associated with offspring birth weight (9). However, high heterogeneity in the meta-analyses, dependence among some of the included studies, and reporting bias may have biased their results. We collected IPD and single-study estimates

from eligible studies in order to pool effects across all studies in a meta-analysis. This approach allowed for exploration of confounding factors and evaluation of preplanned subgroup effects.

METHODS

The systematic review and meta-analysis was reported according to the PRISMA and MOOSE guidelines (11,12), and the protocol was registered at PROSPERO (13).

Study inclusion criteria

We included studies that assessed the association between maternal B12 in serum or plasma during pregnancy and birth weight or gestational age at delivery.

Only studies of longitudinal cohort design were eligible for this review. To be eligible, information on birth weight had to be registered at birth and could not be retrospectively reported and length of gestation, in completed days or weeks, had to be estimated by either ultrasound or last menstrual period, or a combination of the two. Studies where B12 was measured after conception and prior to delivery were eligible. If a study was designed to evaluate women or offspring with a specific condition (e.g. preeclampsia or congenital malformations), and there was a marked overrepresentation of participants with such a condition, that study was excluded. Studies with fewer than 50 participants were not considered. Given the need to collaborate with authors of the original studies, we included only those studies published in 1998 or later.

Search methods

The electronic literature search was constructed by the first author (TR) and a librarian trained in medical database searches, and conducted in PubMed, Scopus, Web of Knowledge, EBSCO-host (CINAHL) and OvidSP (MEDLINE, EMBASE and GLOBAL HEALTH); last accessed August 2015. No language restriction was applied. The reference lists of all studies read in full-text were hand searched to find additional eligible studies. Web Appendix 1 provides complete information on the electronic searches.

Data collection

Electronic literature searches were carried out by the first author (TR). Duplicates were removed and eligibility of all references evaluated by screening of the titles and abstracts by the first author (TR). All potentially eligible studies were read in full-text and assessed for inclusion independently by two authors (TR and KRR). A handsearch of reference lists was done independently by two authors (TR and KRR or MJT). When multiple reports from the same study were found, we used the most complete report.

Risk of bias was independently assessed by two authors (TR and MJT) based on a modified version of the Newcastle-Ottawa Scale (range 0-7) (14). Disagreements were resolved by consulting a third reviewer (KRR). We defined high risk of bias as a score of four or less, and moderate to low risk was defined as scores five through seven.

Authors from all eligible studies were contacted to obtain IPD, each research group being approached at least three times. IPD was received without personal identification. For studies where IPD could not be shared, authors were asked to provide results from pre-specified reanalyzes of their data. When neither IPD nor reanalyzes could be retrieved, relevant estimates were extracted from the publications.

Variables

The main exposure of interest was vitamin B12 levels in maternal serum or plasma. We calculated trimester-specific standard deviation (SD) scores based on studies providing IPD and reanalyzed aggregate data. Analyses were performed for B12-deficiency pre-defined as <148 pmol/L,(15) and B12 tertiles constructed on the basis of included individual data; <148 pmol/L (tertile 1), 148-216 pmol/L (tertile 2), and >216 pmol/L (tertile 3).

The three pre-defined main outcomes were: birth weight as a continuous measure in grams, LBW (birth weight <2,500 g) and small-for-gestational-age (SGA; birth weight SD score <10th centile) (1). Birth weight SD score was calculated using gestational age at delivery and sex-specific reference standards published by the INTERGROWTH 21st Project (16). We assumed birth weight SD score to serve as a proxy of fetal growth, and defined SGA as a proxy of restricted fetal growth. Outcomes related to length of gestation were gestational age at delivery (days) and preterm birth (gestational age at delivery <37 weeks).

Three main confounders were identified based on *a priori* assumptions of confounding factors, availability of data and exploration of effect of covariates on outcome and exposure: maternal age (continuous), pre-pregnancy or pregnancy body mass index (BMI, continuous) and parity (nulliparous versus primiparous or multiparous). Maternal weight was used when information on BMI was unavailable. Also, we considered smoking habits (smoking versus not smoking during pregnancy) and highest completed education (completed high school, equal to 13 years of education, versus not completed high school).

Statistical analysis

We applied a two-step IPD meta-analysis with random effects to pool the results across studies, including aggregate data from individual studies when IPD was not available. All presented results are adjusted for maternal age, BMI/weight and parity (the "main model"), unless otherwise specified. Precision was assessed by 95% confidence intervals (CI).

Mean difference of the continuous outcomes birth weight (g), gestational age at delivery (days) and birth weight SD score (SD) were analyzed by linear regression. To estimate risk ratios (RR), Poisson regression with robust error variance (17) was used to analyze the dichotomous outcomes LBW, SGA, and preterm birth.

We conducted a meta-analysis that evaluated how B12 was associated with maternal weight. Publication bias was explored using funnel plots. Heterogeneity between the studies was explored by computing the I^2 statistic, and was considered to be present when I^2 was greater than 30%. All statistical analyses were carried out using Stata SE version 13.1 (Stata Corporation, College Station, TX, USA). The statistical analyses, including sensitivity analyses, are described in more detail in Web Appendix 2.

RESULTS

Availability of data

The electronic literature search and hand search of reference lists identified 606 unique references (Figure 1). Twenty-two studies met eligibility criteria (11,993 observations) of which 18 studies were included in the meta-analyses (11,216 observations), representing 94% of all eligible observations (18–35). Four eligible studies (777 observations) were not included as they neither reported on the association between maternal B12 and birth weight or length of gestation, nor provided the necessary IPD or results from requested reanalyzes (36–

39). Fourteen of the included studies reported estimates for the association between B12 in pregnancy and birth weight or length of gestation, and were qualitatively appraised in the systematic review section (10,563 observations) (18,19,21,23–25,27,29–35).

For the meta-analyses, ten studies provided IPD (8,928 observations) (18,19,21–23,26–29,32), two studies provided results from reanalyzes (973 observations) (20,35), and relevant information and estimates were extracted from the published reports of six studies (1,315 observations) where IPD or reanalyzes of original data were not provided (24,25,30,31,33,34).

Details of eligible studies

Studies included in the meta-analyses are described in Table 1; details of the eligible studies not included are presented in Web Table 1 (36–39). Of the included studies, one was conducted in North America (34), nine in Europe (18,19,22,25–28,31,32), one in Africa (30), one in Oceania (24), and six in Asia (20,21,23,29,33,35). The number of pregnancies studied ranged from 84 to 5,641. B12 was measured during the first trimester in seven studies (19,22,23,28,31–33), during the second trimester in 15 studies (18–24,26–29,31–33,35), and during the third trimester in 12 studies (18,20,21,23,25,27,29,30,32–35). Mean ± SD B12 concentrations in the first, second and third trimester were 219.8 ± 128.2, 187.8 ± 91.3 and 188.7 ± 82.5 pmol/L, respectively. Preterm deliveries were excluded from four studies (25,26,31,33).

Key maternal and newborn characteristics of the included studies are presented in Table 2. B12-deficiency was identified in 0% to 69% of pregnancies (median 33%). The incidence of LBW ranged from 0% to 33% (median 6%), preterm births from 4% to 14% (median 8%),

and SGA from 5% to 32% (median 11%). Higher maternal weight was associated with lower maternal B12; one SD higher maternal BMI or weight was associated with an 11 pmol/L decrease in B12 (95% CI -15, -7).

Systematic review

Birth weight/SGA. The association between B12 and birth weight or risk of SGA birth was reported in 14 of 22 eligible studies. Three studies reported a clear association: one study reported that birth weight was higher among B12-deficient women than among non-deficient women (34); another study reported that only among women with gestational diabetes mellitus, lower B12 was associated with higher birth weight (32). Conversely, a third study reported that lower values of B12 significantly increased the risk of SGA births (23). In the remaining 11 studies, there was weak evidence of an inverse association in three studies (25,27,33), and no association in eight studies (18,19,21,24,29–31,35).

Length of gestation. Only two published reports reported on the association between B12 and length of gestation or preterm birth. The first study observed that higher B12 was associated with a longer length of gestation and a reduced risk of preterm birth, but the small sample size yielded low precision of the estimates (21). The second study did not find evidence of an association between B12 and length of gestation (19).

Evaluation of the risk of bias showed that the scores ranged between three and seven, and that two studies were classified with high risk of bias (see Web Table 2).

Meta-analysis of maternal B12 in relation to birth weight and LBW

In the meta-analysis, we found no evidence of a linear association between B12 and birth weight (Figure 2): The adjusted estimate was 5.1 g increase in birth weight per SD increase in B12 (95% CI -10.9, 21.0; I^2=30%).

Subgroup and sensitivity analyses are presented in Web Table 3. Stratification by country income showed that there was an association between B12 and birth weight in low- and middle-income countries, but not in high-income countries. Heterogeneity among the studies was explained largely by country income level and maternal BMI or weight. Excluding a study that used late-pregnancy BMI (29), instead of pre-pregnancy or early pregnancy BMI/weight in the other studies, reduced the heterogeneity from I^2=30% to I^2=13% (results not presented). One study reported an association between B12 and birth weight that greatly deviated from the other studies (33). Excluding this study did not notably change the effect estimate, but resulted in a modest reduction in heterogeneity (from I^2=30% to I^2=21%; results not presented). Sensitivity analyses excluding each of the included studies one by one, and excluding studies only evaluating newborns born at term, did not meaningfully alter the association between B12 and birth weight (results not presented).

Results for categories of B12 supported our main results. Neither B12-deficiency nor B12 tertiles were associated with birth weight (see Web Table 4).

B12-deficiency was associated with a 15% (95% CI 1%, 31%; I^2=5%) increased risk of LBW (Figure 3, left panel).

The funnel plot of B12 and birth weight indicated low risk of publication bias (see Web Figure 1).

Since birth weight may be regarded as a summary measure of fetal growth and gestational age, we further performed analyses to assess a possible influence of B12 on these factors.

Meta-analysis of maternal B12 in relation to length of gestation and preterm birth

The analyses did not support a linear association between maternal B12 levels with length of gestation in days (0.1 days (95% CI -0.2, 0.3; I^2=0%) per SD increase of B12). However, *increasing* levels of B12 were associated with a reduced risk of preterm birth (RR 0.89 (95% CI 0.82, 0.97; I^2=0%) per SD increase in B12; Web Figure 2). Accordingly, B12-deficiency in pregnancy was associated with a 21% increased risk of preterm birth (95% CI -1%, 49%; I^2=20%); Figure 3, middle panel).

The association between B12 and preterm birth was similar within all subgroup and sensitivity analyses, although there was a loss of precision in these subgroup analyses due to smaller sample sizes (see Web Table 5).

Meta-analysis of maternal B12 in relation to birth weight SD score and SGA

B12 was not associated with birth weight SD scores in the main analysis (see Web Figure 3). However, B12 was associated with birth weight SD score in low- and middle-income countries (0.08 SD per 1 SD increase in B12 (95% CI 0.03, 0.14; I^2=0%)), but not in high-income countries (-0.02 SD (95% CI -0.05, 0.02; I^2=23%)).

Women with B12-deficiency were not at higher risk of SGA births than non-deficient women (Figure 3, right panel), and B12 levels were similar in SGA and non-SGA pregnancies (see Web Table 6).

DISCUSSION

The results from this systematic review and meta-analysis do not support any linear association between vitamin B12 levels in pregnancy and offspring birth weight. However, our findings provide evidence that lower maternal B12 levels are associated with increased risk of preterm birth, and that the risk of preterm birth was particularly high in the presence of B12-deficiency during pregnancy.

Strengths and limitations

A strength of this study is the use of IPD and reanalyzed data. Due to substantial heterogeneity in the published analyses, a traditional meta-analysis could not answer our research questions. Incomplete or selective reporting may reduce the replicability of studies and distort the literature (40). This is illustrated by comparing the findings of this review with those of a recently published systematic review by Sukumar et al. with traditional meta-analysis on the association between B12 and birth weight (9). That study reported an odds ratio of 1.70 (95% CI 1.16, 2.50; I^2=84%) of the association between "low B12" and "adverse birth weight". A more moderate association was found in the present study in a comparable analysis of B12-deficiency in relation to LBW (RR 1.15 (95% CI 1.01, 1.31); I^2=5%). One reason for the discrepant results may be that Sukumar et al. depended solely on data presented in the published reports and were unable to include results reported as being "insignificant"; for instance, from the largest individual study in the present review (19). The comparable meta-analysis in the present review included roughly ten times as many pregnancies as the meta-analysis in Sukumar's review. Additionally, of eight individual results included in Sukumar's meta-analysis, five evaluated most of the same women from a single original study, exaggerating the influence of a single, outlying study (8,23). By collecting IPD and

requesting reanalyzes from contributing studies, we were able to standardize the analyses across most of the included studies, thereby reducing heterogeneity and facilitating interpretation of results. Compared with the review by Sukumar et al. that presented meta-analyses with high levels of heterogeneity (I^2-scores from 74% to 98% in the primary analyses), the present study had I^2-scores between 0% and 30% in the primary analyses. Additionally, the present study enabled conduction of subgroup and sensitivity analyses, along with more complete adjustment for important confounders (e.g. maternal weight).

We included 94% of all eligible participants, permitting an unbiased summary of the published literature. Given the relative large number of included subjects, we had increased power to evaluate findings reported with low precision in individual studies. We tested the stability of our findings with a broad range of sensitivity analyses.

Another strength was that our analyses were not post-hoc, but followed a detailed protocol. We performed a thorough literature search without language restrictions, and systematically reviewed all eligible studies.

There are several limitations. Unpublished studies were not considered for this review, which could potentially skew estimates. However, a funnel plot did not suggest publication bias. We were unable to include four eligible studies (777 observations, 6% of all observations). Given the small number of observations, it is unlikely that inclusion of these remaining studies would have importantly influenced our main results.

Our approximations of fetal growth and restricted fetal growth by use of gestational age and sex specific birth weight charts is suboptimal, as these outcomes are ideally estimated using

serial ultrasound measurements during pregnancy (41). Furthermore, we did not have sufficient data at hand to evaluate the possible implications of low levels of B12 during different periods in pregnancy in the same woman. Sensitivity analyses stratified by trimester of B12 measurement across studies, however, did not reveal important variation in the association between B12 and the outcomes of interest.

Importantly, B12-deficiency may be a proxy for inadequate nutritional status, and it is possible that some of our findings are related to nutritional status, not specifically to B12. A predominantly plant-based diet is low in B12, but also other nutrients, such as vitamin D and zinc, which to some degree may be associated with preterm birth (42–44). We did not have information on dietary intake or blood levels of these nutrients. Nutritional status could explain the present finding of an association between B12 and birth weight in low- and middle-income countries but not high-income countries. However, lower vitamin B12 levels were associated with higher risk of preterm birth irrespective of country income. It seems less likely that nutritional status can fully explain this finding.

Mixing of effects is inherent in observational studies, and residual confounding cannot be ruled out. We emphasize that our study report associations, and that causal effects must be explored through trials (see below). Reassuringly, we found little discrepancy in the pooled results of adjusted main models as compared to extended adjusted models (i.e. additionally adjusting for maternal education and smoking habits).

Possible explanation of findings

Low birth weight is a result of preterm birth, of being born small at term, or a combination of the two (45). While we found an increased risk of preterm birth and LBW among B12-

deficient women, there was little evidence that maternal B12 levels influenced offspring birth weight SD score or SGA status. It seems more likely that the observed higher risk for LBW in B12-deficient women can be explained by preterm birth rather than by reduced fetal growth.

Higher B12 was associated with higher birth weight in low- and middle-income countries, but not in high-income countries. Four of the five studies included in the low- and middle-income group were performed in an Indian population. Therefore, generalization of these results to low- or middle-income countries outside India should be treated with caution. Indian women generally have lower dietary intake of B12, due to a mainly vegetarian diet, making them susceptible to B12-deficiency (46). Additionally, Indian newborns are among the smallest in the world (45). Our findings suggest that pregnancies already at greatest risk of giving birth to small newborns were the ones most vulnerable to low levels of B12.

The association between B12 and the risk of preterm birth was consistent across studies in both high-income and low- and middle-income countries, and generalization to countries not studied may be feasible.

In line with our findings, maternal obesity has been associated with B12-deficiency in several populations (47,48). It is hypothesized that this association is due to altered fat distribution and metabolism in the overweight compared with normal weight (47). Maternal weight is positively correlated with newborn weight (49), and failure to adjust for maternal weight may underestimate a positive association between B12 and birth weight.

Potential mechanism of action

Preterm birth may be categorized into spontaneous and medically indicated, with varying etiologies (50). Unfortunately, information on spontaneous versus medically indicated preterm births were not available to us. Medically indicated preterm birth are most commonly caused by severe preeclampsia or severely restricted fetal growth (51). Our findings do not support maternal level of B12 to be associated with fetal growth. Maternal B12 may, however, be associated with risk of preeclampsia, potentially through homocysteine, but reports are discrepant (52–54). The rate of medically indicated preterm births is higher in high-income countries than in low- and middle-income countries (55). In analysis stratified by country income, we found similar associations between B12 and risk of preterm birth in low-, middle- and high-income countries. Still, this finding does not link B12 to specific etiologies of preterm birth, which is a topic that deserves further studies.

It is possible that supplementation of B12 or folic acid, with a subsequent reduction of homocysteine, increases birth weight and length of gestation. However, a Cochrane review concluded that supplementation of folic acid during pregnancy did not reduce risk of either preterm birth or LBW (60). Two small (256 pregnancies and 68 pregnancies) randomized controlled trials of B12 supplementation during pregnancy reported on birth weight and length of gestation (61,62). Both observed higher B12 plasma levels in the supplemented group compared with the control group, but no reduction in homocysteine levels. No differences were observed in birth weight, length of gestation, or frequency of LBW births or preterm births in the supplemented group compared with the control group in either study (C Duggan, Harvard University, personal communication, 2015) (61,62). However, the studies were not powered to detect small but meaningful differences in preterm birth.

Context

There are 15 million preterm births and 20 million low birth weight births globally each year (1). The greatest burden of LBW is found in South Asia, while preterm birth is highest in Africa (1). Preterm birth is the leading cause of neonatal deaths (1). In the era of The Millennium Development Goals (1990-2015), post-neonatal under-five mortality rate was reduced by 58% (2). Reduction in neonatal mortality was less pronounced (47%) (2). Prevention of preterm birth is thus a key strategy to reduce neonatal deaths and reach the new target of under-five year mortality of 25 per 1,000 live births by 2030, down from 43 per 1,000 in 2015 (2).

Our systematic review was not designed to study the prevalence of B12-deficiency during pregnancy. However, this condition was common in the studies in our review, and comparable to a systematic review of B12-deficiency during pregnancy (9). A large group of women are thus affected by a potential preventable risk of preterm birth.

Conclusion and implications for clinical practice and future research

Vitamin B12-deficiency during pregnancy is common. Results of this systematic review with IPD meta-analyses provides robust evidence that lower B12 levels during pregnancy are associated with increased risk of preterm birth, particularly in B12-deficient women. Our findings support conducting randomized controlled trials to evaluate whether maternal B12 supplementation in pregnancy reduces the risk of preterm birth.

Acknowledgement

We would like to acknowledge the following persons for their help in order to provide the necessary data for the present study: Professor Caroline H D Fall, MRC Lifecourse Epidemiology Unit, University of Southampton, UK; Dr Simon Wheeler, nutritionist at the MRC Epidemiology Unit, University of Cambridge; Professor Anura V Kurpad, Professor of Physiology, St. John's Medical College, India; and Dr Nithya Sukumar, research fellow at Warwick Medical School, University of Warwick, UK. Also thanks to Professor Christopher Duggan, Professor in the Department of Nutrition, Harvard University, USA, for providing necessary unpublished results from a previous trial. Lastly, we would like to acknowledge Janis Glover, Academic Librarian at the Medical Library, Yale University, USA, who helped develop the search strategy.

References

1. Lawn JE, Blencowe H, Oza S, et al. Every Newborn: progress, priorities, and potential beyond survival. *Lancet*. 2014;384(9938):189–205.
2. UNICEF, WHO, World Bank, et al. Levels and trends in child mortality 2015. 2015 1-36 p.
3. Barros FC, Bhutta Z a, Batra M, et al. Global report on preterm birth and stillbirth (3 of 7): evidence for effectiveness of interventions. *BMC Pregnancy Childbirth*. 2010;10(Suppl 1):S3.
4. Allen LH. Vitamin B12 metabolism and status during pregnancy, lactation and infancy. *Adv. Exp. Med. Biol.* 1994;352:173–86.
5. Allen RH, Stabler SP, Savage DG, et al. Metabolic abnormalities in cobalamin (vitamin B12) and folate deficiency. *FASEB J.* 1993;7(14):1344–53.
6. Chiang PK, Gordon RK, Tal J, et al. S-Adenosylmethionine and methylation. *FASEB J.* 1996;10(4):471–80.
7. Halarnkar PP, Blomquist GJ. Comparative aspects of propionate metabolism. *Comp. Biochem. Physiol. B.* 1989;92(2):227–31.
8. Muthayya S, Kurpad A, Duggan C, et al. Low maternal vitamin B12 status is associated with intrauterine growth retardation in urban South Indians. *Eur. J. Clin. Nutr.* 2006;60(6):791–801.
9. Sukumar N, Rafnsson S, Kandala N, et al. Prevalence of vitamin B-12 insufficiency during pregnancy and its effect on offspring birth weight: a systematic review and meta-analysis. *Am. J. Clin. Nutr.* 2016;103(5):1232–51.
10. Haider BA, Bhutta ZA. Multiple-micronutrient supplementation for women during pregnancy. *Cochrane database Syst. Rev.* 2015;11:CD004905.
11. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*. 2009;339:b2700.
12. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA*. 2000;283(15):2008–12.
13. University of York. PROSPERO. 2015;(http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42015025141). (Accessed January 15, 2016)
14. Wells G, Shea B, O'Connell D, et al. The Newcastle-Ottawa Scale for assessing the quality of nonrandomised studies in meta-analyses. 2013;(http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp)
15. Stabler SP. Vitamin B12 Deficiency. *N. Engl. J. Med.* 2013;368(2):149–160.
16. Villar J, Ismail LC, Victora CG, et al. International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. *Lancet*. 2014;384:857–868.
17. Zou G. A modified poisson regression approach to prospective studies with binary

data. *Am J Epidemiol*. 2004;159(7):702–706.

18. Baker PN, Wheeler SJ, Sanders TA, et al. A prospective study of micronutrient status in adolescent pregnancy. *Am. J. Clin. Nutr.* 2009;89(4):1114–1124.

19. Bergen NE, Jaddoe VW V, Timmermans S, et al. Homocysteine and folate concentrations in early pregnancy and the risk of adverse pregnancy outcomes: The Generation R Study. *BJOG*. 2012;119(6):739–51.

20. Bhate VK, Joshi SM, Ladkat RS, et al. Vitamin B12 and folate during pregnancy and offspring motor, mental and social development at 2 years of age. *J. Dev. Orig. Health Dis.* 2012;3(2):123–130.

21. Chen LW, Lim AL, Colega M, et al. Maternal folate status, but not that of vitamins B-12 or B-6, is associated with gestational age and preterm birth risk in a multiethnic Asian population. *J. Nutr.* 2015;145(1):113–20.

22. Dayaldasani A, Ruiz-Escalera J, Rodriguez-Espinosa M, et al. Serum vitamin B12 levels during the first trimester of pregnancy correlate with newborn screening markers of vitamin B12 deficiency. *Int. J. Vitam. Nutr. Res.* 2014;84(1-2):92–7.

23. Dwarkanath P, Barzilay JR, Thomas T, et al. High folate and low vitamin B-12 intakes during pregnancy are associated with small-for-gestational age infants in South Indian women: a prospective observational cohort study. *Am. J. Clin. Nutr.* 2013;98(6):1450–8.

24. Furness D, Fenech M, Dekker G, et al. Folate, vitamin B12, vitamin B6 and homocysteine: impact on pregnancy outcome. *Matern. Child Nutr.* 2013;9(2):155–66.

25. Halicioglu O, Sutcuoglu S, Koc F, et al. Vitamin B12 and folate statuses are associated with diet in pregnant women, but not with anthropometric measurements in term newborns. *J. Matern. Fetal. Neonatal Med.* 2012;25(9):1618–21.

26. Hay G, Clausen T, Whitelaw A, et al. Maternal folate and cobalamin status predicts vitamin status in newborns and 6-month-old infants. *J Nutr*. 2010;140(3):557–64.

27. Hogeveen M, Blom HJ, van der Heijden EH, et al. Maternal homocysteine and related B vitamins as risk factors for low birthweight. *Am. J. Obstet. Gynecol.* 2010;202(6):572.e1–6.

28. Kaymaz C, Demir A, Bige O, et al. Analysis of perinatal outcome by combination of first trimester maternal plasma homocysteine with uterine artery Doppler velocimetry. *Prenat. Diagn.* 2011;31(13):1246–50.

29. Krishnaveni G V., Veena SR, Karat SC, et al. Association between maternal folate concentrations during pregnancy and insulin resistance in Indian children. *Diabetologia*. 2014;57(1):110–21.

30. Mamabolo RL, Alberts M, Steyn NP, et al. The effect of maternal glucose metabolism, iron, vitamin B12 and folate status on pregnancy outcomes. *South African J. Clin. Nutr.* 2006;19(3):120–30.

31. Relton CL, Pearce MS, Parker L. The influence of erythrocyte folate and serum vitamin B12 status on birth weight. *Br. J. Nutr.* 2005;93(05):593.

32. Sukumar N, Bawazeer N, Patel V, et al. Low B12 Level Is Associated with Maternal Obesity and Higher Birthweight in Gestational Diabetes. In: *Journal of Developmental Origins of Health and Disease*. 2011:128–9.

33. Takimoto H, Mito N, Umegaki K, et al. Relationship between dietary folate intakes, maternal plasma total homocysteine and B-vitamins during pregnancy and fetal growth in Japan. *Eur. J. Nutr.* 2007;46(5):300–6.

34. Wu BTF, Innis SM, Mulder K a., et al. Low plasma vitamin B-12 is associated with a lower pregnancy-associated rise in plasma free choline in Canadian pregnant women and lower postnatal growth rates in their male infants. *Am. J. Clin. Nutr.* 2013;98(5):1209–1217.

35. Yajnik CS, Deshpande SS, Jackson a a, et al. Vitamin B12 and folate concentrations during pregnancy and insulin resistance in the offspring: the Pune Maternal Nutrition Study. *Diabetologia*. 2008;51(1):29–38.

36. Karakantza M, Androutsopoulos G, Mougiou A, et al. Inheritance and perinatal consequences of inherited thrombophilia in Greece. *Int. J. Gynecol. Obstet.* 2008;100(2):124–129.

37. López-Quesada E, Vilaseca MA, Vela A, et al. Perinatal outcome prediction by maternal homocysteine and uterine artery Doppler velocimetry. *Eur. J. Obstet. Gynecol. Reprod. Biol.* 2004;113(1):61–66.

38. Lee S, Guillet R, Cooper EM, et al. Maternal Inflammation at Delivery Affects Assessment of Maternal Iron Status. *J. Nutr.* 2014;144(10):1524–32.

39. Neumann CG, Oace SM, Chaparro MP, et al. Low vitamin B12 intake during pregnancy and lactation and low breastmilk vitamin 12 content in rural Kenyan women consuming predominantly maize diets. *Food Nutr Bull*. 2013;34(2):151–9.

40. Glasziou P, Altman DG, Bossuyt P, et al. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet*. 2014;383(9913):267–276.

41. Nyberg D a, Abuhamad A, Ville Y. Ultrasound assessment of abnormal fetal growth. *Semin. Perinatol.* 2004;28(1):3–22.

42. Position of the American Dietetic Association and Dietitians of Canada: Vegetarian diets. *J. Am. Diet. Assoc.* 2003;103(6):748–65.

43. Ota E, Mori R, Middleton P, et al. Zinc supplementation for improving pregnancy and infant outcome. *Cochrane database Syst. Rev.* 2015;2:CD000230.

44. De-Regil LM, Palacios C, Lombardo LK, et al. Vitamin D supplementation for women during pregnancy. *Cochrane database Syst. Rev.* 2016;1:CD008873.

45. United Nations Children's Fund and World Health Organization. Low Birthweight: Country, regional and global estimates. 2004 1-31 p.

46. Yajnik CS, Chandak GR, Joglekar C, et al. Maternal homocysteine in pregnancy and offspring birthweight: epidemiological associations and Mendelian randomization analysis. *Int. J. Epidemiol.* 2014;43(5):1487–97.

47. Knight BA, Shields BM, Brook A, et al. Lower Circulating B12 Is Associated with Higher Obesity and Insulin Resistance during Pregnancy in a Non-Diabetic White British Population. *PLoS One*. 2015;10(8):e0135268.

48. Krishnaveni G V., Hill JC, Veena SR, et al. Low plasma vitamin B12 in pregnancy is associated with gestational "diabesity" and later diabetes. *Diabetologia*. 2009;52(11):2350–2358.

49. Romundstad PR, Smith GD, Nilsen TIL, et al. Associations of prepregnancy cardiovascular risk factors with the offspring's birth weight. *Am. J. Epidemiol.* 2007;166(12):1359–1364.

50. Goldenberg RL, Culhane JF, Iams JD, et al. Epidemiology and causes of preterm birth. *Lancet*. 2008;371(9606):75–84.

51. Gyamfi-Bannerman C, Fuchs KM, Young OM, et al. Nonspontaneous late preterm birth: Etiology and outcomes. *Am. J. Obstet. Gynecol.* 2011;205(5):456.e1–456.e6.

52. Wadhwani NS, Patil V V, Mehendale SS, et al. Increased homocysteine levels exist in women with preeclampsia from early pregnancy. *J. Matern. Fetal. Neonatal Med.* 2016;29(16):2719–25.

53. Makedos G, Papanicolaou A, Hitoglou A, et al. Homocysteine, folic acid and B12 serum levels in pregnancy complicated with preeclampsia. *Arch. Gynecol. Obstet.* 2007;275(2):121–124.

54. Sanchez SE, Zhang C, Malinow MR, et al. Plasma folate, vitamin B12, and homocyst(e)ine concentrations in preeclamptic and normotensive Peruvian women. *Am. J. Epidemiol.* 2001;153(5):474–480.

55. Blencowe H, Cousens S, Oestergaard MZ, et al. National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. *Lancet*. 2012;379(9832):2162–72.

56. Hogeveen M, Blom HJ, den Heijer M. Maternal homocysteine and small-for-gestational-age offspring: systematic review and meta-analysis. *Am. J. Clin. Nutr.* 2012;95(1):130–6.

57. Miner SES, Evrovski J, Cole DEC. Clinical chemistry and molecular biology of homocysteine metabolism: An update. *Clin. Biochem.* 1997;30(3):189–201.

58. Refsum H, Ueland PM, Nygård O, et al. Homocysteine and cardiovascular disease. *Annu. Rev. Med.* 1998;49(1):31–62.

59. Dodds L, Fell DB, Dooley KC, et al. Effect of homocysteine concentration in early pregnancy on gestational hypertensive disorders and other pregnancy outcomes. *Clin. Chem.* 2008;54(2):326–334.

60. Lassi ZS, Salam RA, Haider BA, et al. Folic acid supplementation during pregnancy for maternal health and pregnancy outcomes. *Cochrane database Syst. Rev.* 2013;3:CD006896.

61. Duggan C, Srinivasan K, Thomas T, et al. Vitamin B-12 supplementation during pregnancy and early lactation increases maternal, breast milk, and infant measures of vitamin B-12 status. *J. Nutr.* 2014;144(5):758–64.

62. Siddiqua TJ, Ahmad SM, Ahsan KB, et al. Vitamin B12 supplementation during pregnancy and postpartum improves B12 status of both mothers and infants but vaccine response in mothers only: a randomized clinical trial in Bangladesh. *Eur. J. Nutr.* 2016;55(1):281–93.

Legends to figures

Figure 1 *Title:* Flow chart of studies included in at least one of the meta-analyses of the association between B12 and birth weight or length of gestation

Legend: B12, vitamin B12; IPD, individual participant data; N, number of studies; n, number of pregnancies.
a, IPD or reanalyses not provided, and results could not be abstracted from published reports.

Figure 2 *Title:* Forest plot presenting the association between B12 and birth weight

Legend: Meta-analysis of studies of the association between vitamin B12 and birth weight after adjustment for maternal age, parity and body mass index or weight. Effect estimates are expressed as change in birth weight per one standard deviation increase of vitamin B12. CI, confidence interval; n, number pregnancies.

Figure 3 *Title:* Forest plot presenting the association between B12-deficiency and the risk of low birth weight (left panel), preterm birth (middle panel), and small-for-gestational-age birth (right panel)

Legend: Meta-analysis of studies of the association between vitamin B12-deficiency and the risk of low birth weight (left panel), preterm birth (middle panel), and small-for-gestational-age birth (right panel) after adjustment for maternal age, parity and body mass index or weight. Effect estimate expressed as risk ratio of the outcome comparing B12 deficient to non-deficient. CI, confidence interval; n, number pregnancies; RR, risk ratio.

Web Figure 1 *Title:* Funnel plot of studies evaluating the association between B12 and birth weight

Legend: Funnel plot of studies evaluating the association between vitamin B12 and birth weight after adjustment for maternal age, parity and body mass index or weight. Individual studies are represented by solid dots, and the pseudo-95% confidence interval by broken lines.

Web Figure 2 *Title:* Forest plot presenting the association between B12 and the risk of preterm birth

Legend: Meta-analysis of studies of the association between vitamin B12 and the risk of preterm birth after adjustment for maternal age, parity and body mass index or weight. Effect estimates are expressed as risk ratios of preterm birth per one standard deviation increase of vitamin B12. CI, confidence interval; n, number pregnancies; RR, risk ratio.

Web Figure 3 *Title:* Forest plot presenting the association between B12 and birth weight SD score

Legend: Meta-analysis of studies of the association between vitamin B12 and birth weight standard deviation scores (i.e. accounting for length of gestation and sex) after adjustment for maternal age, parity and body mass index or weight. Effect estimate expressed as change in birth weight standard deviation score per one standard deviation increase of vitamin B12. CI, confidence interval; n, number pregnancies.

Table 1. Characteristics of Studies Included in the Meta-Analysis

						Week of B12 measurement		Included in specific meta-analyses[a]					
Study	**Data**	**n**	**Country**	**Study years**	**B12 analysis method**	**Range**	**Median**	**Birth weight**	**LBW**	**SGA**	**Birth weight SD score**	**Length of gestation**	**Preterm birth**
Baker, 2009 (18)	IPD	290	The United Kingdom	2004-2007	RIA	27-43	30	x	x	x	x	x	x
Bergen, 2012 (19)	IPD	5,641	The Netherlands	2002-2006	ECL	5-18	13	x	x	x	x	x	x
Bhate, 2012 (20)	Reanalysed data	214	India	2004-2006	Micro-biological	24-30	28	x	x			x	x
Chen, 2015 (21)	IPD	988	Singapore	2009-2010	ECL	26-29	27	x	x	x	x	x	x
Dayaldasani, 2014 (22)	IPD	187	Spain	2011	ECL	3-23	10	x[b]	x[b]	x[b]	x[b]	x[b]	x[b]
Dwarkanath, 2013 (23)	IPD	344	India	2001-2003	ECL	T1: 5-19 T2: 20-29 T3: 30-39	T1: 12 T2: 24 T3: 34	x	x	x	x	x	x
Furness, 2013 (24)	Data from publication	84	Australia	NA	ECL	18-20	NA			x[c]			
Halicioglu, 2012 (25)	Data from publication	208	Turkey	2008	ECL	>37	NA	x[d]					
Hay, 2010 (26)	IPD	149	Norway	1997	Micro-biological	17-19	NA	x					
Hogeveen, 2010 (27)	IPD	363	The Netherlands	2002-2004	Micro-biological	27-38	31	x[e]	x[e]	x[e]	x[e]	x[e]	x[e]
Kaymaz, 2011 (28)	IPD	103	Turkey	2007	ECL	11-14	13	x	x			x	x

Table 1. Continued

Krishnaveni, 2014 (29)	IPD	654	India	1997-1998	Micro-biological	22-35	26	x	x	x	x	x	x
Mamabolo, 2006 (30)	Data from publication	219	South Africa	1999-2000	RIA	28-36	NA			x[c]			
Relton, 2005 (31)	Data from publication	500	The United Kingdom	2000-2002	RIA	NA	11.5 (5.8)[f]				x[e]		
Sukumar, 2011 (32)	IPD	209	The United Kingdom	2005-2010	RIA (n=182), ECL (n=27)	0-37	24	x	x	x	x	x	x
Takimoto, 2007 (33)	Data from publication	88	Japan	2001-2003	ECL	T1: 7-14 T3: 34-36	T1: NA T3: NA	x[g]					
Wu, 2013 (34)	Data from publication	216	Canada	NA	RIA	NA	36	x[d]					
Yajnik, 2008 (35)	Reanalysed data	759	India	1994-1996	Micro-biological	NA	T2: 18 (2)[f]	x	x			x	x

Studies are referred to according to their citation number in the text. ECL, electroluminescence; IPD, individual participant data; n, number of pregnancies; NA, not available; RIA, radioimmunoassay; SD, standard deviation; SGA, small-for-gestational-age; T1, 1st trimester; T2, 2nd trimester; T3, 3rd trimester.

a, included in the analyses of the exposures B12 SD score and B12-deficiency, both crude and adjusted (maternal age, body mass index or weight, and parity), if not otherwise specified;

b, does not contribute in the analyses of B12-deficiency (none of the participants were deficient);

c, level of B12 in SGA versus non-SGA, crude analysis;

d, birth weight among B12-deficient versus non-deficient, crude analysis;

e, crude analysis;

f, mean (SD);

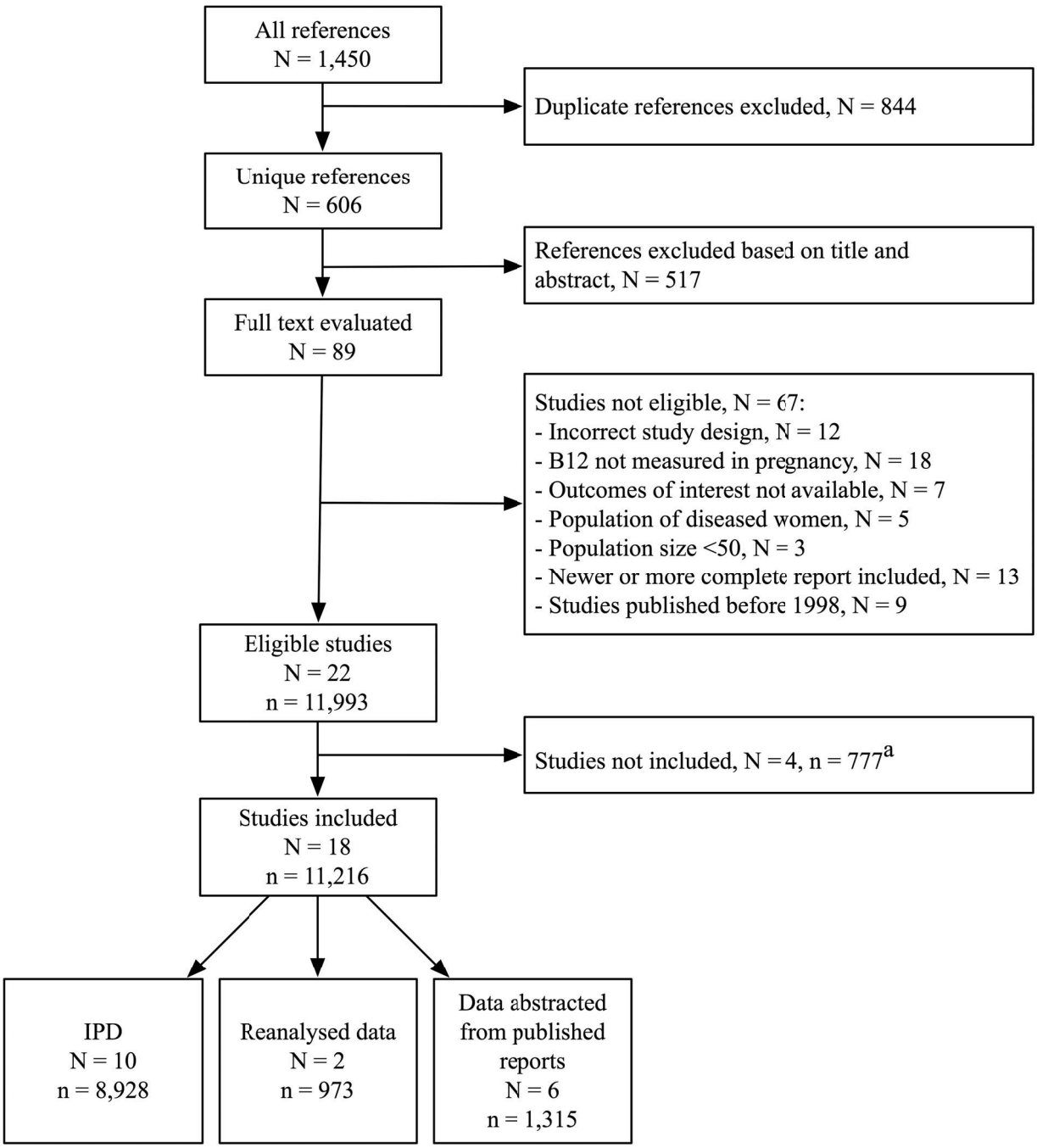
g, adjusted analysis (maternal age, body mass index or weight, and parity).

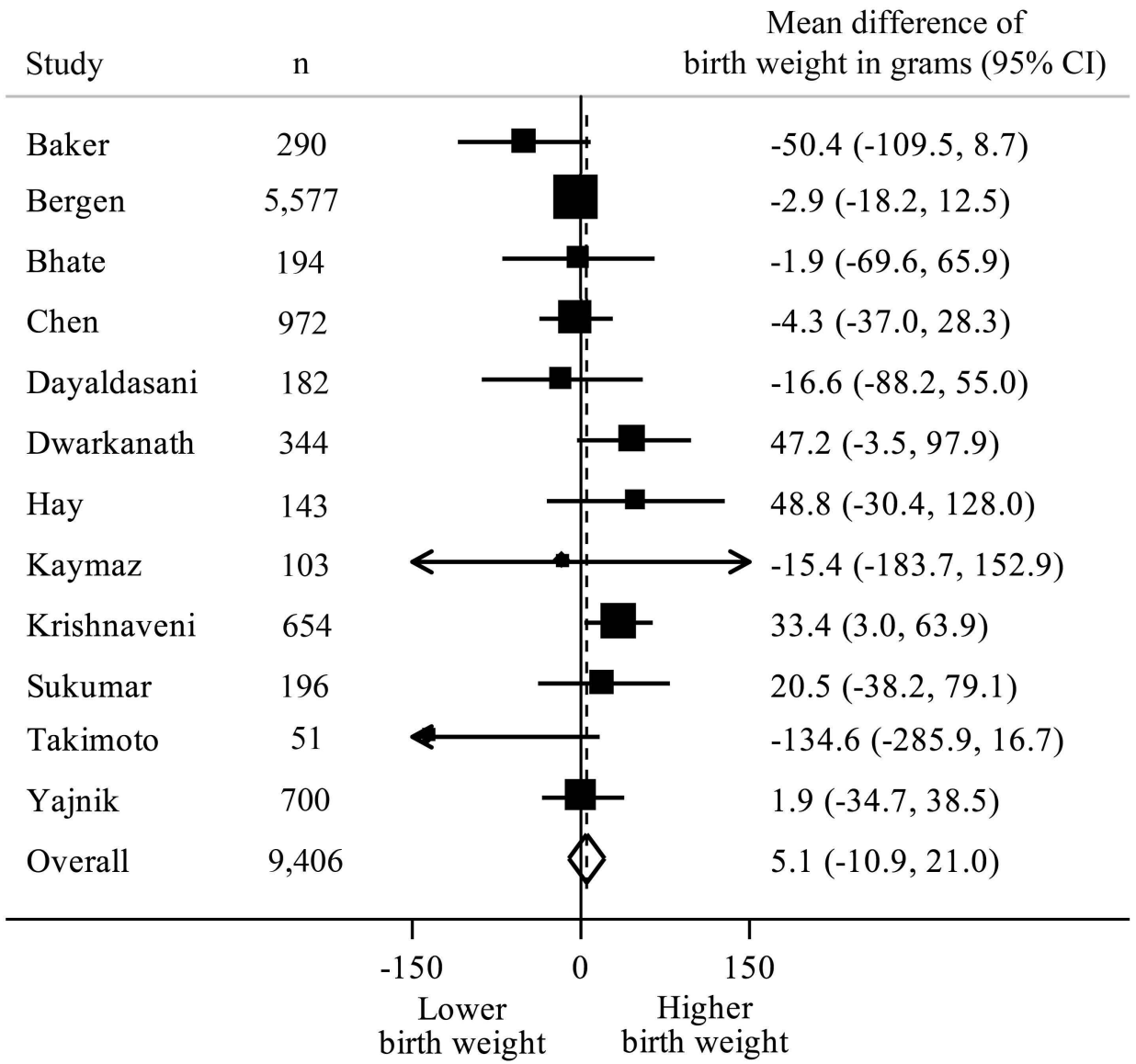
Table 2. Maternal and Newborn Characteristics of Studies Included in the Meta-Analysis

Study	Maternal age (years), mean (SD)	Maternal BMI[o] (kg/m^2), mean (SD)	Para 0		Vitamin B12 (pmol/L), mean (SD)	B12-deficient[a]		Birth weight (g), mean (SD)	LBW[b]		SGA[c]		Length of gestation (weeks), mean (SD)	Preterm birth[d]	
			n	%		n	%		n	%	n	%		n	%
Baker, 2009 (18)	18 (1)	65 (14)[e]	277	96	192 (84)	93	32	3,232 (534)	26	9	33	12	39.7 (1.8)	22	8
Bergen, 2012 (19)	30 (5)	25 (5)	3,208	57	188 (93)	2,098	37	3,418 (563)	280	5	412	7	39.9 (1.8)	268	5
Bhate, 2012 (20)	23 (3)	20 (3)	165	71	145 (84)	148	69	2,707 (411)	49	25	NA	NA	38.6 (2.6)	18	8
Chen, 2015 (21)	31 (5)	66 (12)[e]	420	43	220 (79)	161	16	3,101 (449)	76	8	86	9	38.6 (1.4)	85	9
Dayaldasani, 2014 (22)	30 (6)	26 (5)	96	51	387 (123)	0	0	3,267 (526)	11	6	12	7	38.8 (1.9)	14	8
Dwarkanath, 2013 (23)	24 (4)	53 (10)[e]	203	59	205 (115)[f]	100	29[f]	2,771 (498)	95	28	102	30	38.3 (1.7)	47	14
Furness, 2013[g] (24)	33 (7)	27 (5)	NA	NA	234 (129)	NA	NA	3,390 (789)	NA	NA	21	25[h]	38.8 (2.9)	NA	NA
Halicioglu, 2012[g] (25)	28 (5)	NA	NA	NA	120[i]	99	48[j]	3,357 (466)	NA	NA	NA	NA	NA	NA	NA
Hay, 2010 (26)	30 (4)	65 (10)[e]	67	45	294 (87)	2	1	3,727 (476)	0	0	NA	NA	NA	NA	NA
Hogeveen, 2010 (27)	33 (4)	NA	109	30	186 (69)	120	34	3,436 (545)	18	5	19	5	39.5 (1.6)	21	6
Kaymaz, 2011 (28)	27 (3)	24 (4)	45	44	152 (59)	54	52	3,241 (553)	5	5	NA	NA	38.4 (1.9)	9	9
Krishnaveni, 2013 (29)	24 (4)	24 (4)	331	51	187 (100)	264	40	2,857 (475)	126	19	202	32	39.0 (1.8)	63	10
Mamabolo, 2006[g] (30)	25 (7)	27 (4)	NA	NA	175 (77)	36	16[k]	3,120 (550)	NA	NA	66	30[l]	NA	NA	NA
Relton, 2005[g] (31)	28 (6)[m]	NA	NA	43[m]	239 (97)	NA	NA	3,430 (470)[m]	NA	NA	NA	NA	NA	NA	NA
Sukumar, 2011 (32)	31 (6)	27 (6)	68	33	168 (126)	114	55	3,381 (558)	10	5	16	8	39.3 (1.7)	9	4
Takimoto, 2007[g] (33)	29 (5)	21 (3)	NA	NA	405 (146)[f]	13	16[n]	3,120 (411)	5	5	NA	NA	39.6 (1.0)	NA	NA
Wu, 2013[g] (34)	33 (4)	NA	NA	NA	224 (96)	51	24	3,486 (452)	NA	NA	NA	NA	NA	NA	NA
Yajnik, 2008 (35)	21 (4)	18 (2)	252	31	151 (78)	447	59	2,612 (392)	230	33	NA	NA	38.8 (2.1)	87	11

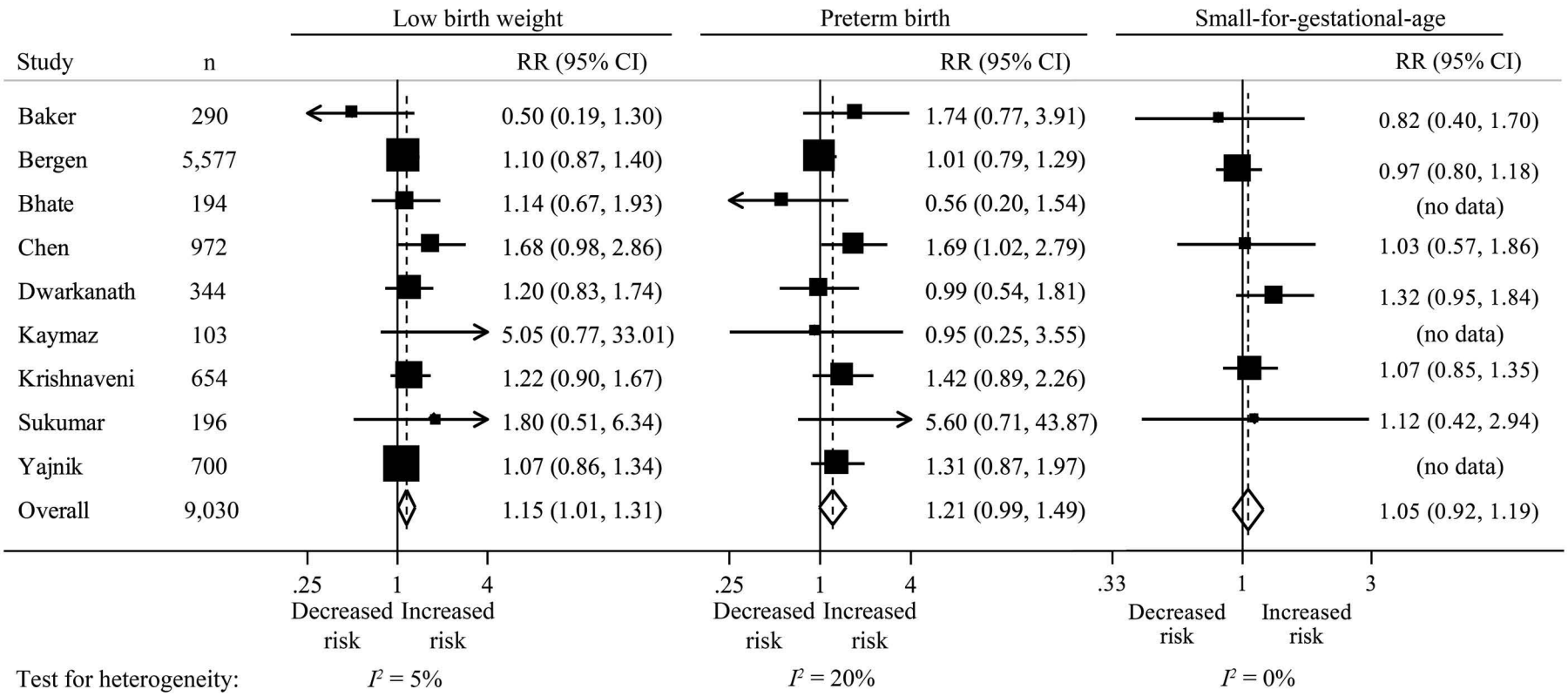
Studies are referred to according to their citation number in the text. BMI, body mass index; LBW, low birth weight; NA, not available; SD, standard deviaton; SGA, small-for-gestational-age.

a, B12 <148 pmol/L;
b, birth weight <2,500 g;
c, birth weight SD score (i.e. accounting for length of gestation and sex) below 10th centile;
d, length of gestation <37 weeks;
e, kg (BMI not available);
f, first measurement;
g, data extracted from publication;
h, serial tapering of growth in abdominal circumference and of estimated fetal weight below the 10th centile of an Australian growth chart;
i, median (range not available);
j, B12 ≤118 pmol/L;
k, B12-deficiency not defined;
l, lowest birth weight tertile (mean birth weight 2,940 g) used as approximation of SGA for the purpose of this review;
m, based on a larger study population than the subgroup with available B12 data included in this review (n=974-997);
n, third trimester
o, weight (kg)/height (m)^2.


All references
N = 1,450
Duplicate references excluded, N = 844
Unique references
N = 606
References excluded based on title and abstract, N = 517
Full text evaluated
N = 89
Studies not eligible, N = 67:
- Incorrect study design, N = 12
- B12 not measured in pregnancy, N = 18
- Outcomes of interest not available, N = 7
- Population of diseased women, N = 5
- Population size <50, N = 3
- Newer or more complete report included, N = 13
- Studies published before 1998, N = 9
Eligible studies
N = 22
n = 11,993
Studies not included, N = 4, n = 777[a]
Studies included
N = 18
n = 11,216
IPD
N = 10
n = 8,928
Reanalysed data
N = 2
n = 973
Data abstracted from published reports
N = 6
n = 1,315



Test for heterogeneity: $I^2 = 30\%$


Low birth weight
Preterm birth
Small-for-gestational-age
Study
n
RR (95% CI)
RR (95% CI)
RR (95% CI)
Baker 290 0.50 (0.19, 1.30) 1.74 (0.77, 3.91) 0.82 (0.40, 1.70)
Bergen 5,577 1.10 (0.87, 1.40) 1.01 (0.79, 1.29) 0.97 (0.80, 1.18)
Bhate 194 1.14 (0.67, 1.93) 0.56 (0.20, 1.54) (no data)
Chen 972 1.68 (0.98, 2.86) 1.69 (1.02, 2.79) 1.03 (0.57, 1.86)
Dwarkanath 344 1.20 (0.83, 1.74) 0.99 (0.54, 1.81) 1.32 (0.95, 1.84)
Kaymaz 103 5.05 (0.77, 33.01) 0.95 (0.25, 3.55) (no data)
Krishnaveni 654 1.22 (0.90, 1.67) 1.42 (0.89, 2.26) 1.07 (0.85, 1.35)
Sukumar 196 1.80 (0.51, 6.34) 5.60 (0.71, 43.87) 1.12 (0.42, 2.94)
Yajnik 700 1.07 (0.86, 1.34) 1.31 (0.87, 1.97) (no data)
Overall 9,030 1.15 (1.01, 1.31) 1.21 (0.99, 1.49) 1.05 (0.92, 1.19)
.25
1
4
Decreased risk
Increased risk
.25
1
4
Decreased risk
Increased risk
.33
1
3
Decreased risk
Increased risk
Test for heterogeneity:
I^2 = 5%
I^2 = 20%
I^2 = 0%

Web Appendix 1. Search Terms.

The search terms were adapted to each service provider and database, and were composed of a combination of the following (or related) terms: B12 and pregnancy and birth weight or length of gestation. We added restriction terms excluding review articles, intervention studies and case reports, studies evaluating adults, children (other than infants), rodents, and patients with anemia. We used a combination of controlled vocabulary terms and free text words.

Pubmed

("Vitamin B 12"[Mesh] OR B12[Text Word] OR "B 12"[Text Word] OR cobalamin*[Text Word]) AND (pregnan*[Text Word] OR Pregnancy[Mesh] OR gestation*[Text Word] OR fetus[MeSH] OR fetus*[Text Word] OR foetus*[Text Word] OR foetal* [Text Word] OR fetal*[Text Word] OR ”Fetal Development”[Mesh] OR ”Infant, Newborn”[Mesh]) AND ("Infant, Low Birth Weight"[Mesh] OR "Birth Weight"[Mesh] OR birthweight[Text Word] OR "birth weight"[Text Word] OR SGA[Text Word] OR "fetal growth retardation"[MeSH] OR IUGR[Text Word] OR ”growth restriction”[Text Word] OR ”growth retardation”[Text Word] OR "small for gestational age"[Text Word] OR "small for date"[Text Word] OR ”Infant, Premature”[Mesh] OR ”Premature Birth”[Mesh] OR ”Gestational Age”[Mesh] OR preterm[Text Word] OR prematur*[Text Word] OR ”gestational age”[Text Word] OR ”length of gestation”[Text Word] OR ”duration of pregnancy”[Text Word]) NOT (”Review”[Publication Type] OR ”Child”[Mesh] OR ”Aged”[Mesh] OR ”Case Reports”[Publication Type] OR ”Clinical Trial”[Publication Type] OR ”Rodentia”[Mesh] OR ”Anemia”[Mesh])

OvidSP Medline

(exp Vitamin B 12/ or b12.tw or B 12.tw or cobalamin*.tw) and (pregnan*.tw or exp Pregnancy/ or gestation*.tw or exp Fetus/ or fetus*.tw or fetal*.tw or foetus*.tw or foetal*.tw or exp Fetal Development/ or exp Infant, Newborn/) and (exp Infant, Low Birth Weight/ or exp Birth Weight/ or birth weight.tw or birthweight.tw or SGA.tw or exp Fetal Growth Retardation/ or IUGR.tw or growth restriction.tw or growth retardation.tw or small for gestational age.tw or small for date.tw or exp Infant, Premature/ or exp Premature Birth/ or exp Gestational Age/ or preterm.tw or prematur*.tw or gestational age.tw or length of gestation.tw or duration of pregnancy.tw) not (review/ or exp child/ or exp aged/ or exp case report/ or exp clinical trial/ or exp rodentia/ or exp anemia/)

OvidSP Embase

(exp cyanocobalamin/ or exp cyanocobalamin deficiency/ or exp cobalamin derivative/ or exp cobalamin/ or b12.tw or b 12.tw or cobalamin*.tw) and (pregnan*.tw or exp pregnancy/ or gestation*.tw or exp fetus/ or fetus*.tw or

fetal*.tw or foetus*.tw or foetal*.tw or exp fetus growth/ or exp newborn/) and (exp birth weight/ or birthweight.tw or birth weight.tw or SGA.tw or exp intrauterine growth retardation/ or IUGR.tw or growth restriction.tw or growth retardation.tw or small for gestational age.tw or small for date.tw or exp prematurity/ or exp premature labor/ or exp gestational age/ or preterm.tw or premature.tw or gestational age.tw or length of gestation.tw or duration of pregnancy.tw) not (exp review/ or exp case report/ or exp aged/ or exp anemia/ or exp clinical trial/ or exp rodent/)

OvidSP Global Health

(exp vitamin b12/ or b12.tw or b 12.tw or cobalamin*.tw) and (pregnan*.tw or exp pregnancy/ or gestation*.tw or exp fetus/ or fetus*.tw or fetal*.tw or foetus*.tw or foetal*.tw or exp fetal development/ or exp neonates/) and (exp low birth weight infants/ or exp birth weight/ or birthweight.tw or birth weight.tw or sga.tw or exp growth retardation/ or iugr.tw or growth restriction.tw or growth retardation.tw or small for gestational age.tw or small for date.tw or exp prematurity/ or exp premature infants/ or exp gestation period/ or preterm.tw or prematur*.tw or gestational age.tw or length of gestation.tw or duration of pregnancy.tw) not (exp reviews/ or exp elderly/ or exp case reports/ or exp clinical trials/ or exp rodents/ or anaemia.sh)

EBSCO-host CINAHL

((MH ”Vitamin B 12") OR (MH ”Vitamin B12 Deficiency+”) OR b12 OR "b 12" OR cobalamin*) AND (pregnan* OR (MH ”Pregnancy+”) OR gestation* OR (MH ”Fetus+”) OR fetus* OR foetus* OR foetal* OR fetal* OR (MH ”Infant, Newborn+”)) AND ((MH ”Infant, Low Birth Weight+") OR (MH "Birth Weight") OR birthweight OR "birth weight" OR SGA OR (MH "Fetal Growth Retardation") OR IUGR OR ”growth restriction” OR ”growth retardation” OR "small for gestational age" OR "small for date" OR (MH ”Infant, Premature”) OR (MH ”Childbirth, Premature”) OR (MH ”Gestational Age”) OR preterm OR prematur* OR ”gestational age” OR ”length of gestation” OR ”duration of pregnancy”) NOT ((MH ”Literature Review+”) OR (MH ”Child, Preschool”) OR (MH ”Aged+”) OR (MH ”Case Studies”) OR (MH ”Clinical Trials+”) OR (MH ”Rodents+”) OR (MH ”Anemia+”))

SCOPUS

(TITLE-ABS-KEY(b12 OR ”b 12” OR cobalamin*)) AND (TITLE-ABS-KEY(pregnan* OR gestation* OR fetus* OR fetal* OR foetus* OR foetal* OR newborn*)) AND (TITLE-ABS-KEY(”birth weight” OR ”birthweight” OR sga OR ”growth retardation” OR ”growth restriction” OR iugr OR ”small for gestational age” OR ”small for date” OR preterm OR prematur* OR ”gestational age” OR ”length of gestation” OR ”duration of pregnancy”)) AND (EXCLUDE(DOCTYPE, ”re”)) AND NOT ((TITLE(anemi* OR anaemi*)) OR (TITLE-ABS-KEY(mouse OR mice OR rat OR rats OR rodent*)))

Web of Knowledge

#1: (TS=(b12 OR ”b 12” OR cobalamin*)) AND (TS=(pregnan* OR gestation* OR fetus* OR fetal* OR foetus* OR foetal* OR newborn*)) AND (TS=(”birth weight” OR ”birthweight” OR sga OR ”growth retardation” OR ”growth restriction” OR iugr OR ”small for gestational age” OR ”small for date” OR preterm OR prematur* OR ”gestational age” OR ”length of gestation” OR ”duration of pregnancy”)) NOT ((TI=(anemi* OR anaemi*)) OR (TS=(mouse OR mice OR ra OR rats OR rodent*)))

#2: Restrict #1 to reviews

#3: #1 NOT #2

Web Appendix 2. Statistical Analyses.

A multivariable model was applied adjusting for maternal age, BMI (or weight when BMI was unavailable) and parity. When IPD was not provided, we requested results from the following reanalyzes of original studies: the association of B12 (SD score) with birth weight, gestational age at delivery, LBW and preterm birth; and the association of B12-deficiency with birth weight, LBW and preterm birth. Results were provided for both crude analyses, and two different multivariable analyses (adjusting for maternal age, BMI and parity; and adjusting for maternal age, BMI, parity and smoking habits). When neither IPD nor results from requested reanalyzes were available, we extracted relevant results from the publications.

We stratified our analysis for the following *a priori* subgroup and sensitivity analyses: trimester of B12 measurement (four strata: 1st, 2nd, 3rd trimesters, and 1st and 2nd trimesters combined), country income category (high-income versus low- and middle-income countries, as defined by The World Bank),[1] risk of bias (high risk versus moderate or low risk of bias), and excluding each of the studies one by one. Additional sensitivity analyses that were carried out: overweight status (BMI $\geq$25 kg/m^2 versus BMI<25 kg/m^2), B12 assay technique (radioimmunoassay, electroluminescence, microbiological), alternative multivariable models (e.g. a more saturated model including maternal education and smoking habits in addition to the main model), fixed effects model, Poisson regression with non-robust error variance, logistic regression model (dichotomous outcomes), and by excluding studies that only evaluated newborns born at term.

References:

1. The World Bank. The World Bank. (2016). at <http://data.worldbank.org/country>

Risk of bias scale

1) Was B12 ascertained irrespective of the risk of LBW or PTB, and otherwise not prone to selection bias?
2) The study controlled for maternal BMI or weight either by matching or by statistical methods?
3) The study controlled for previous LBW births or PTB, or maternal age, parity, socioeconomic status (SES), smoking habits, ethnicity, vegetarian status or B12 supplement use (at least two of these) either by matching or by statistical methods? *(Because of potential over-adjustment, if a study adjusted for levels of folate, homocysteine or methylmalonic acid, they earned no point on this item (even if they had adjusted for two or more of the mentioned confounders))*
4) Was the exposed cohort truly or somewhat representative of the average pregnant population in the community?
5) Did the women with B12-deficiency receive the same follow-up and interventions as the non-deficient women? *(e.g. not similar if start of multivitamin supplementation if B12-deficient but not non-deficient)*
6) Was the outcome assessed by independent or blind assessment, or by secure records or record linkage?
7) >80% follow-up or description provided for those lost to follow-up?

Web Table 1. Study Characteristics of Eligible Studies Not Included in the Meta-Analysis[a]

Study	n	Country	Study years	B12 analysis method	Week of B12 measurement, range, median	Main objectives
Karakantza, 2008 (36)	392	Greece	2004-2006	NA	6-8	Studied three thrombophilic mutations in relation to pregnancy outcomes (including IUGR).
Lee, 2014 (38)	T2: 83 T3: 42	USA	2006-2012	RIA	T2: 24.4 ± 2.2[b] T3: 29.7 ± 1.8[b]	Studied changes in iron status in pregnant adolescents, and iron status in relation to hepcidin and inflammatory markers.
López-Quesada, 2004 (37)	94	Spain	2000-2001	ECL	24-24, median 24	Studied uterine artery Doppler velocimetry in relation to pregnancy outcomes and homocysteine, folate and B12.
Neumann, 2013 (39)	138	Kenya	1984-1986	NA	NA	Studied B12 dietary intake during pregnancy and lactation in relation to pregnancy outcome, breast milk B12 concentration and infant growth and development.

Studies are referred to according to their citation number in the text. ECL, electroluminescence; IUGR, intrauterine growth restriction; n, number of pregnancies; NA, not available; RIA, radioimmunoassay; T2, 2nd trimester; T3, 3rd trimester.

a, Eligible studies were not included in the systematic review or the meta-analyses when individual participant data or results from requested reanalyses were not provided, appropriate data and results were not available in the original report, and when no association between B12 and birth weight or length of gestation was presented;

b, mean ± SD.

Web Table 2. Risk of Bias of Studies Included in the Meta-Analysis

Study	1	2	3	4	5	6	7	Total
Baker, 2009 (18)	1	1	1	0	1	1	0	5
Bergen, 2012 (19)	1	1	1	1	1	1	1	7
Bhate, 2012 (20)	1	1	1	1	1	1	1	7
Chen, 2015 (21)	1	1	1	1	1	1	1	7
Dayaldasani, 2014 (22)	1	1	1	1	1	?	1	6
Dwarkanath, 2013 (23)	1	1	1	1	1	1	0	6
Furness, 2013 (24)	1	1	1	0	1	1	1	6
Halicioglu, 2012 (25)	1	0	0	1	1	1	1	5
Hay, 2010 (26)	1	1	1	0	1	1	0	5
Hogeveen, 2010 (27)	1	0	1	1	1	1	0	5
Kaymaz, 2011 (28)	1	1	1	1	1	1	?	6
Krishnaveni, 2013 (29)	1	1	1	1	1	1	0	6
Mamabolo, 2006 (30)	1	0	0	1	1	1	1	5
Relton, 2005 (31)	1	0	0	1	?	1	0	3
Sukumar, 2011 (32)	1	1	1	1	1	1	0	6
Takimoto, 2007 (33)	1	1	1	1	1	1	0	6
Wu, 2013 (34)	1	0	0	?	1	1	?	3
Yajnik, 2008 (35)	1	1	1	1	1	1	1	7

Each item was scored "1" (i.e. "yes"), "0" (i.e. "no") or "?" (i.e. "uncertain"), where only the answer "1" scored 1 point. The following questions were evaluated: 1: Was B12 ascertained irrespective of the risk of low birth weight birth or preterm birth, and otherwise not prone to selection bias?; 2: Was the study controlled for maternal body mass index or weight either by matching or by statistical methods?; 3: Was the study controlled for previous low birth weight birth or preterm birth, or maternal age, parity, socioeconomic status, smoking habits, ethnicity, vegetarian status or B12 supplement use (at least two of these) either by matching or by statistical methods? In addition, because of potential over-adjustment, if a study adjusted for levels of folate, homocysteine or methylmalonic acid, they earned no point on this item (even if they had adjusted for two or more of the mentioned confounders); 4: Was the exposed cohort truly or somewhat representative of the average pregnant population in the community?; 5: Did the women with B12-deficiency receive the same follow-up and interventions as the non-deficient women?; 6: Was the outcome assessed by independent or blind assessment, or by secure records or record linkage?; 7: Was follow-up >80% or was any description provided for those lost to follow-up? Studies are referred to according to their citation number in the text.

Web Table 3. Pooled Results From Subgroup and Sensitivity Analyses of B12 SD Score on Birth Weight

Analysis	Number of studies	Number of pregnancies	Birth weight (g) per 1 SD increase in B12 (95% CI)	I^2
Alternative models				
Crude	12[18-23,26-29,32,35]	9,819	-4.9 (-15.7, 5.8)	0
Adjusting for BMI or weight[a]	9[18,19,21-23,26,28,29,32]	8,505	7.8 (-9.0, 24.6)	28
Adjusting for maternal age,[b] parity,[c] and BMI or weight[a] ("main model")	12[18-23,26,28,29,32,33,35]	9,406	5.1 (-10.9, 21.0)	30
Adjusting for maternal age,[b] parity,[c] BMI or weight,[a] and smoking[d]	10[18-22,26,28,29,32,35]	8,420	0.7 (-15.5, 16.9)	29
Adjusting for maternal age,[b] parity,[c] BMI or weight,[a] smoking[d] and education[e] ("extended model")	4[18,19,26,29]	5,948	-1.9 (-40.5, 36.6)	76
Adjusted main model among those with data on smoking	10[18-22,26,28,29,32,35]	8,420	2.4 (-8.8, 13.6)	0
Adjusted main model among those with data on the extended model	4[18,19,26,29]	5,948	4.1 (-27.8, 35.9)	65
Adjusting for BMI[b] among those with weight	4[19,21,22,29]	7,416	6.3 (-10.5, 23.1)	26
Adjusting for weight[b] among those with BMI	4[19,21,22,29]	7,416	8.2 (-6.2, 22.5)	13
Fixed effects model	12[18-23,26,28,29,32,33,35]	9,406	3.4 (-7.4, 14.1)	30
Trimester of B12 measurement				
1st trimester	4[19,22,23,32]	1,461	19.1 (-13.1, 51.4)	0
2nd trimester	8[19,21-23,26,28,29,32]	6,217	11.5 (-11.4, 34.5)	34
1st and 2nd trimester	10[19,21-23,26,28,29,32,33,35]	8,325	12.1 (-9.7, 33.8)	47
3rd trimester	6[18,21,23,29,32,33]	1,140	-0.1 (-31.0, 30.8)	14
Measurement technique[f]				
Radioimmunoassay	2[18,32]	459	-16.1 (-84.9, 52.7)	61
Electroluminescence assay	7[19,21-23,28,32,33]	7,256	0.0 (-18.3, 18.4)	14
Microbiologic assay	4[20,26,29,35]	1,691	20.4 (-0.9, 41.7)	0
Country income category				
High income	7[18,19,21,22,26,32,33]	7,411	-5.5 (-24.8, 13.7)	23
Middle or low income	5[20,23,28,29,35]	1,995	22.2 (2.1, 42.4)	0

(continued)

Web Table 3. Continued

Maternal BMI				
BMI <25 kg/m^2	6[19,21,22,28,29,32]	4,728	-1.9 (-16.4, 12.5)	0
BMI ≥25 kg/m^2	6[19,21,22,28,29,32]	2,945	17.5 (-17.6, 52.5)	41

Pooled results of the mean difference (95% CI) in birth weight (g) per 1 SD increase in maternal B12. All analyses are linear regression analyses with random effects, and adjusted for the main model (i.e. maternal age, BMI (weight if missing BMI), and parity (nulliparous yes/no)) unless otherwise specified. Studies included in the analyses are referred to according to their citation number in the text. BMI, body mass index; CI, confidence interval; SD, standard deviation.

a, BMI and weight (if missing BMI) as continuous covariates;

b, continuous covariate;

c, nulliparous (yes/no);

d, smoking during pregnancy (yes/no);

e, completed high school (yes/no);

f, Sukumar 2011[32] measured n=182 by radioimmunoassay and n=27 by electroluminescence assay.

Web Table 4. B12-Deficiency and B12 Tertiles in Relation to Birth Weight

Analysis	Number of studies	Number of pregnancies	Number of exposed	Mean difference in birth weight (g) in exposed versus non-exposed (95% CI)	I^2
B12-deficiency					
IPD, crude	9[18,19,21,23,26-29,32]	8,735	3,006	9.3 (-14.9, 33.5)	0
IPD, main model[a]	8[18,19,21,23,26,28,29,32]	8,279	2,846	-14.5 (-39.1, 10.2)	0
Aggregate, crude[b]	4[20,25,34,35]	1,323	695	63.8 (-32.6, 159.9)	69
Aggregate, main model[a]	2[20,35]	894	542	-1.36 (-53.5, 50.8)	0
IPD + aggregate, crude	13[18-21,23,25-29,32,34,35]	10,058	3,701	23.3 (-6.7, 53.4)	24
IPD + aggregate, main model[a]	10[18-21,23,26,28,29,32,35]	9,173	3,388	-14.0 (-36.3, 8.3)	0
B12 tertiles[c]					
IPD, crude	10[18,19,21-23,26-29,32]	5,942	3,179	18.9 (-15.8, 53.8)	8
IPD, main model[a]	9[18,19,21-23,26,28,29,32]	5,633	2,997	-16.6 (-54.5, 21.2)	11

Pooled results of the mean difference in birth weight (g) in exposed versus non-exposed pregnancies. All analyses are random effects models and crude, unless otherwise specified. Studies included in the analyses are referred to according to their citation number in the text. CI, confidence interval; IPD, individual participant data.

a, adjusted for maternal age, body mass index (weight if missing body mass index), and parity (nulliparous yes/no);

b, B12-deficiency defined as <148 pmol/L except for Halicioglu 2012[25] (<118 pmol/L);

c, lowest tertile (i.e. exposed) versus highest tertile.

Web Table 5. Pooled Results From Subgroup and Sensitivity Analyses of B12 SD Score and the Risk of Preterm Birth

Analysis	Number of studies	Number of pregnancies	Number of preterm births	Risk ratio of preterm birth per 1 SD increase in B12 (95% CI)	I^2
Alternative models					
Crude	11[18-23,27-29,32,35]	9,747	643	0.90 (0.81, 1.00)	15
Adjusting for BMI or weight[a]	8[18,19,21-23,28,29,32]	8,362	510	0.89 (0.81, 0.98)	0
Adjusting for maternal age,[b] parity,[c] and BMI or weight[a] ("main model")	10[18-23,28,29,32,35]	9,291	615	0.89 (0.82, 0.97)	0
Adjusting for maternal age,[b] parity,[c] BMI or weight,[a] and smoking[d]	9[18-22,28,29,32,35]	8,365	531	0.90 (0.82, 0.99)	0
Adjusting for maternal age,[b] parity,[c] BMI or weight,[a] smoking[d] and education[e] ("extended model")	3[18,19,29]	5,813	304	0.92 (0.82, 1.03)	0
Adjusted main model among those with data on smoking	9[18-22,28,29,32,35]	8,365	531	0.89 (0.81, 0.98)	0
Adjusted main model among those with data on the extended model	3[18,19,29]	5,813	304	0.91 (0.81, 1.02)	0
Adjusting for BMI[b] among those with data on weight	4[19-22,29]	7,416	424	0.91 (0.82, 1.01)	0
Adjusting for weight[b] among those with data on BMI	4[19-22,29]	7,416	424	0.90 (0.81, 1.00)	0
Fixed effects model	10[18-23,28,29,32,35]	9,291	615	0.89 (0.82, 0.97)	0
Non-robust error variance	10[18-23,28,29,32,35]	9,291	615	0.90 (0.82, 0.99)	0
Logistic regression	10[18-23,28,29,32,35]	9,285	615	0.88 (0.80, 0.97)[f]	0
Trimester of B12 measurement					
1st trimester	3[19,22,23]	1,453	107	1.01 (0.84, 1.20)	0
2nd trimester	6[19,21,23,28,29,32]	6,061	344	0.85 (0.67, 1.07)	57
1st and 2nd trimesters	8[19,21-23,28,29,32,35]	8,190	522	0.86 (0.71, 1.05)	45
3rd trimester	5[18,21,23,29,32]	1,058	86	0.86 (0.70, 1.06)	4
Measurement technique					
Radioimmunoassay	2[18,32]	459	29	0.60 (0.33, 1.07)	56
Electroluminescence assay	5[19,21-23,28]	7,178	418	0.93 (0.83, 1.04)	0
Microbiologic assay	3[20,29,35]	1,627	168	0.86 (0.73, 1.02)	0
Country income category					
High income	5[18,19,21,22,32]	7,217	391	0.87 (0.74, 1.02)	19
Middle or low income	5[20,23,28,29,35]	2,074	224	0.88 (0.76, 1.02)	0

(continued

Web Table 5. Continued

Maternal BMI					
BMI <25 kg/m^2	6[19,21,22,28,29,32]	4,728	241	0.90 (0.81, 1.01)	0
BMI ≥25 kg/m^2	5[19,21,22,29,32]	2,913	200	0.81 (0.63, 1.05)	44

Pooled results of the risk ratio (95% CI) of preterm birth per 1 SD increase in maternal B12. All analyses are Poisson regression analyses with random effects, robust error variance and adjusted for the main model (maternal age, BMI (weight if missing BMI), and parity (nulliparous yes/no)) unless otherwise specified. Studies included in the analyses are referred to according to their citation number in the text. BMI, body mass index; CI, confidence interval; SD, standard deviation.

a, BMI and weight (if missing BMI) as continuous covariates;

b, continuous covariate;

c, nulliparous (yes/no);

d, smoking during pregnancy (yes/no);

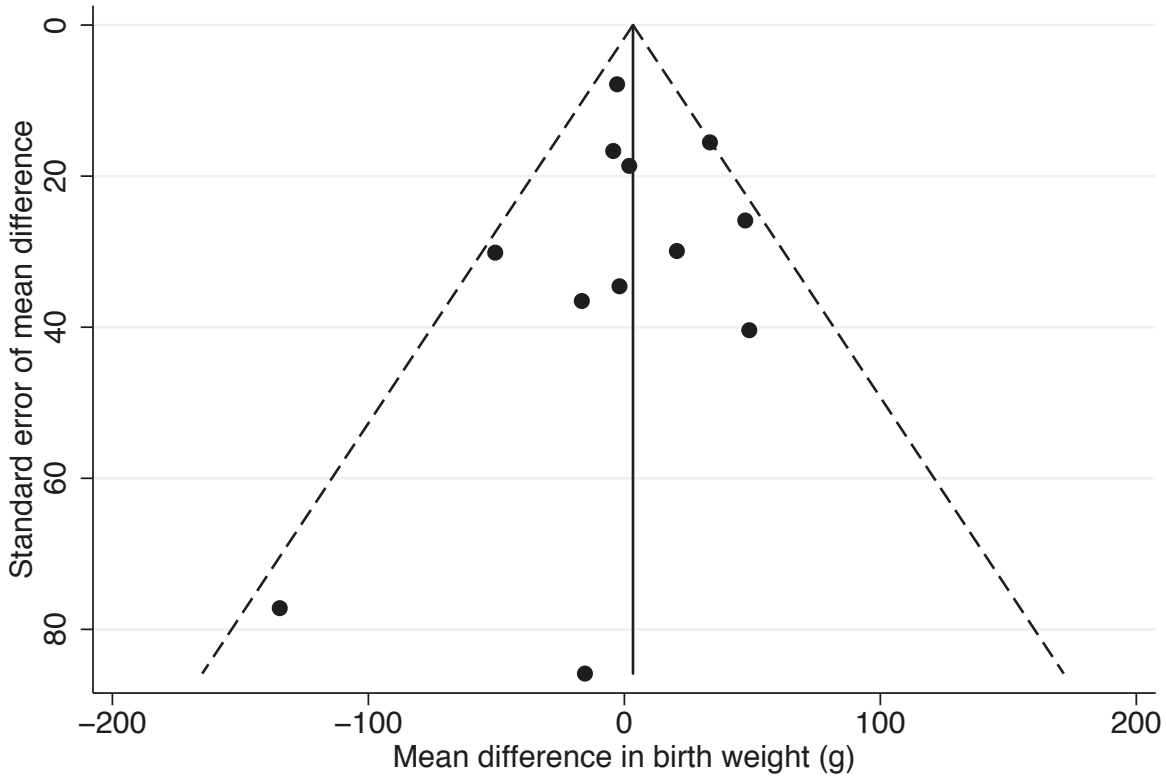
e, completed high school (yes/no);

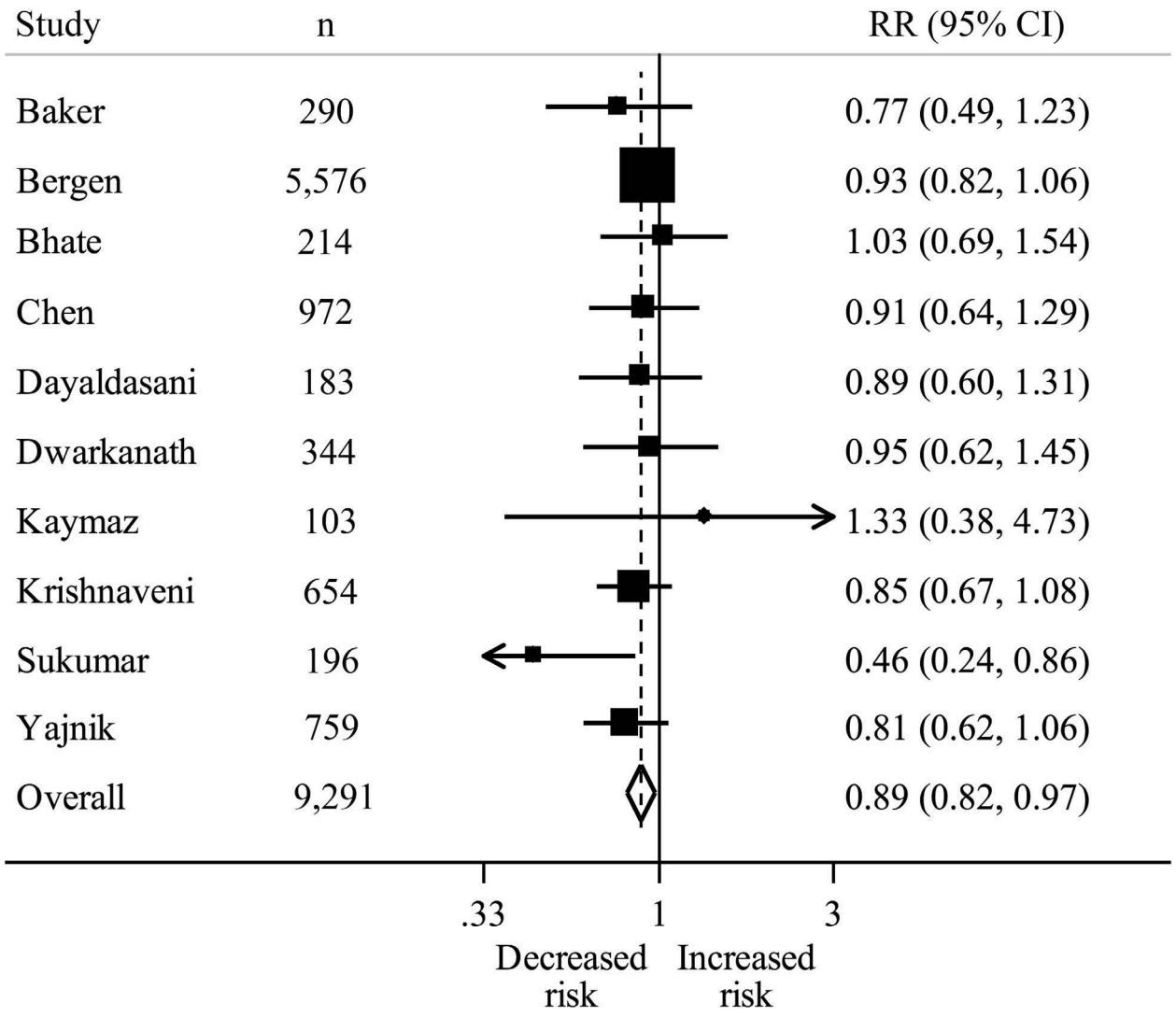
f, odds ratio.

Web Table 6. Level of B12 in Small-for-Gestational-Age and Non-Small-for-Gestational-Age Pregnancies

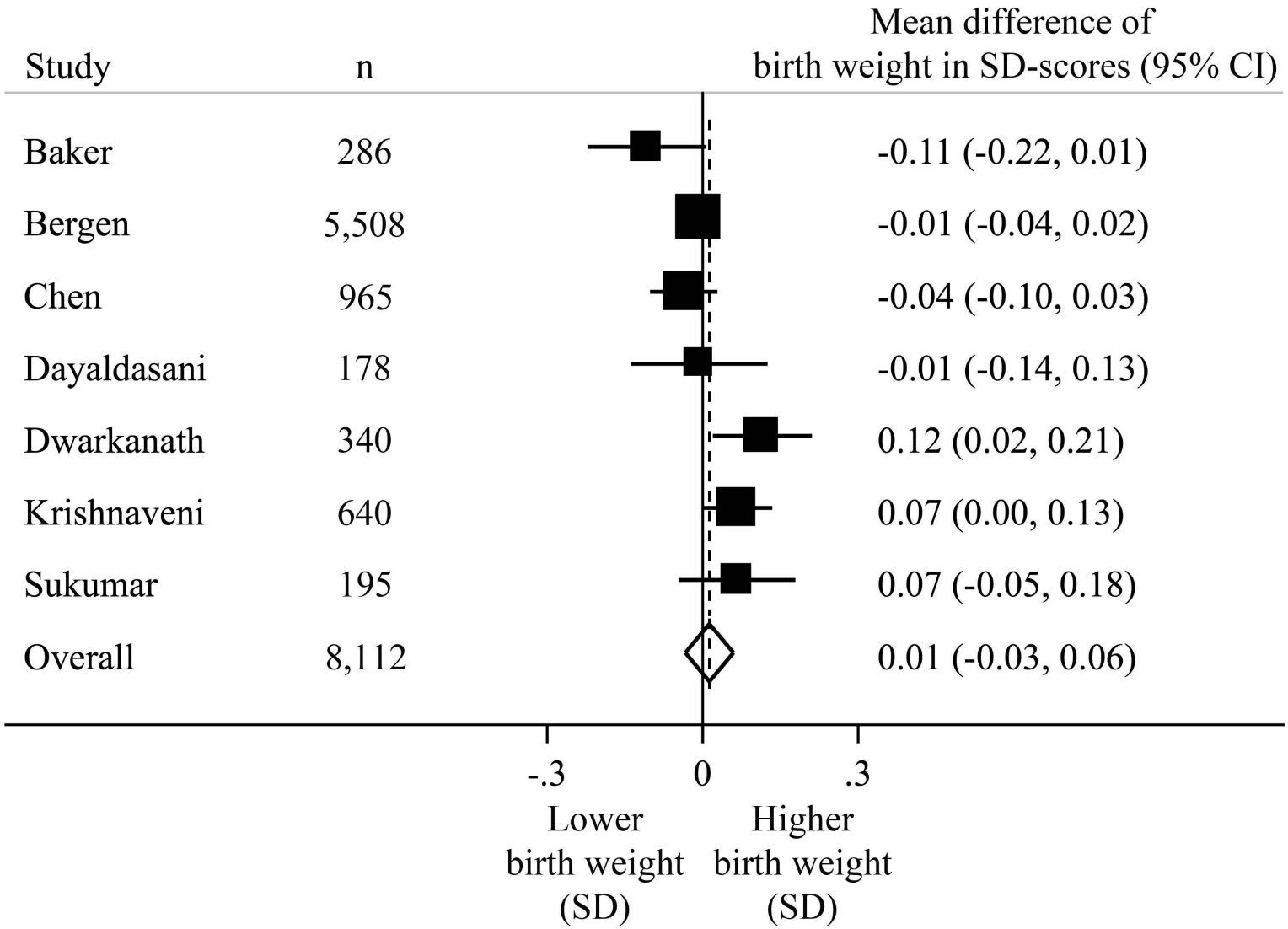
Analysis	Number of studies	Number of pregnancies	Number of SGA[a] births	Mean difference in B12 (pmol/L) in SGA versus non-SGA (95% CI)	I^2
IPD	8[18,19,21-23,27,29,32]	8,561	882	3.3 (-3.4, 9.9)	0
Aggregate	2[24,30]	303	87	-11.7 (-47.4, 24.1)	49
IPD + aggregate	10[18,19,21-24,27,29,30,32]	8,864	969	2.4 (-3.9, 8.7)	0

Pooled results of the mean difference in maternal B12 (pmol/L) in SGA versus non-SGA pregnancies. All analyses are random effects models and crude. Studies included in the analyses are referred to according to their citation number in the text. CI, confidence interval; IPD, individual participant data; SGA, small-for-gestational-age.
a, SGA defined as birth weight <10th centile after taking sex and length of gestation into account, with the following exceptions: In Furness 2013,[24] SGA was defined as "serial tapering of growth in abdominal circumference and of estimated fetal weight below the 10th centile of population-based growth charts", and in Mamabolo 2006,[30] SGA was defined as birth weight in the 1st tertile (mean 2,940 g).


Standard error of mean difference
0
20
40
60
80
−200
−100
0
100
200
Mean difference in birth weight (g)



Test for heterogeneity: $I^2 = 0\%$



Test for heterogeneity: $I^2 = 62\%$