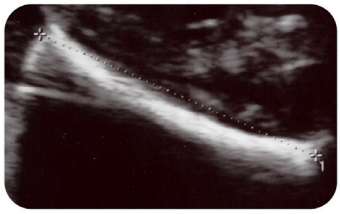
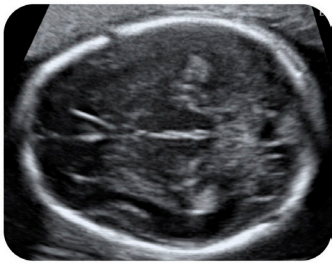
Inger Økland

Biases in second-trimester ultrasound dating related to prediction models and fetal measurements





Thesis for the degree of Doctor Philosophiae

Stavanger, January 2012

Norwegian University of Science and Technology
Faculty of Medicine
Department of Laboratory Medicine,
Children's and Women's Health

NTNU – Trondheim
Norwegian University of
Science and Technology

National Center for
Fetal Medicine
ST. OLAVS HOSPITAL
UNIVERSITETSSYKEHUSET I TRONDHEIM

Stavanger University Hospital
Stavanger Hospital Trust

NTNU
Norwegian University of Science and Technology

Thesis for the degree of Philosophiae Doctor

Faculty of Medicine
Department of Laboratory Medicine, Children's and Women's Health

ISBN 978-82-471-3261-6 (printed ver.)
ISBN 978-82-471-3262-3 (electronic ver.)
ISSN 1503-8181

Doctoral theses at NTNU, 2012:4

Printed by NTNU-trykk

FEILKILDER VED ULTRALYD-BASERT TERMINFASTSETTELSE I 2. TRIMESTER, RELATERT TIL PREDIKSJONSMETODER OG MÅLINGER

Rutineundersøkelsen med ultralyd rundt svangerskapsuke 17–19 har vært en viktig del av norsk svangerskapsomsorg siden den ble innført i 1986. Når skjer fødselen? — altså sikker fastsettelse av fødselstermin, er ett av spørsmålene rutineundersøkelsen skal gi svar på. Terminen er blitt bestemt ved at fosterstørrelsen måles med ultralyd og relateres til antatt fosteralder og derfra til beregnet ultralydtermin. Termindato er vanligvis basert på målinger av fosterets hodestørrelse, men lengden av lårbeinet kan også brukes. Det er viktig at modellene som brukes til terminberegning og målingene de baserer seg på, ikke har svakheter som fører til utilsiktede, systematiske feil i ultralydterminene.

Vi har undersøkt 3 norske modeller for terminbestemmelse, 'Snurra', 'Terminhjulet' og 'eSnurra', og sammenliknet terminen som ble fastsatt med hver modell med det faktiske fødselstidspunktet. Vi har også gjort en studie der vi har vurdert om ultralydmålingene av fosterlårbein endres over tid som følge av tekniske forbedringer i ultralydmaskinene.

Studie 1: I en database med 41 343 rutineundersøkelser utført ved St. Olavs Hospital, Trondheim, sammenliknet vi kvaliteten på terminene som ble fastsatt med de 2 tradisjonelle modellene Snurra og Terminhjulet. Vi fant systematiske avvik; Snurra beregnet termin for seint og Terminhjulet for tidlig i forhold til reelt fødselstidspunkt. Avvikene varierte med fosterstørrelsen på undersøkelsestidspunktet, de strakte seg fra 0 til 4 dager og i hver sin retning. Sannsynligvis skyldes avvikene at det litt snevre datagrunnlaget til de tradisjonelle modellene ikke er godt nok tilpasset den populasjonen de brukes i.

Studie 2: I tillegg til de 2 tradisjonelle modellenes beregninger, så vi nå også på terminene som ble bestemt med en ny populasjonsbasert modell, eSnurra. Vi brukte en database med 9046 rutineundersøkelser fra Stavanger Universitetssjukehus. Resultatene for de 2 første modellene tilsvarte resultatene i Studie 1 både for terminer beregnet fra hodemål og fra lårbeinsmål. eSnurra predikerte stabile, korrekte ultralydterminer, uavhengig av fosterstørrelsen ved undersøkelsen.

Studie 3: Vi ønsket å se om resultatene fra Studie 1 og 2 lot seg reprodusere også i en tredje populasjon og analyserte derfor 23 020 rutineundersøkelse fra Oppland fylke. Også her var resultatene stabile for eSnurra; avviket mellom fødselstidspunkt og ultralydtermin var stort sett mindre enn 1 døgn. For de 2 andre modellene var mønsteret det samme som tidligere påvist. Avvikene virker uunngåelige med de tradisjonelle, seleksjonsbaserte modellene.

Studie 4: Strålebredden i ultralydapparatene blir smalere når teknologien blir bedre. Dette kan tenkes å påvirke lengdemålinger av strukturer som måles på tvers av lydstrålens retning. Vi sammenliknet strålebredde i gamle og nye maskiner og analyserte deretter 41 941 ultralydmålinger av fosterlårbein, samlet over en 18-års periode. Tekniske forbedringer har redusert strålebredden, og dette påvirker enkelte ultralydmålinger ved at strukturen blir målt kortere. Gamle måletabeller kan gi feil resultat.

Konklusjon: En undersøkelse som tilbyes alle gravide bør være standardisert og resultatene til å stole på. En termindato med avvik på 4 dager kan ha konsekvenser både for behandling av svært preterme fostre og for håndtering av overtidige svangerskap. En populasjonsbasert modell for terminfastsettelse (eSnurra) synes å være bedre tilpasset norske svangerskap enn de tradisjonelle modellene.

Kandidat:	***Inger Økland***
Institutt:	***Laboratoriemedisin, barne- og kvinnesykdommer, Norges teknisk-naturvitenskapelige universitet (NTNU)***
Hovedveileder:	***Sturla H. Eik-Nes***
Biveiledere:	***Håkon K. Gjessing og Per Grøttum***

Avhandlingen er funnet verdig til å forsvares offentlig for graden doctor philosophiae i klinisk medisin.
Disputas finner sted i Auditoriet ØHA11, Øya Helsehus, Trondheim, fredag 13. januar 2012 kl. 10.15.

Content

Acknowledgements

Five years ago, my ‘academic journey’ started, and the destination — the completion of this thesis — now seems to be within reach. It has been incredibly inspiring, developing, enlightening, with lots of discussions, laughter and fun. At times it has also been exhausting, frustrating and emotionally challenging, but *never* boring!

To be offered an opportunity of studying term prediction with Professor Sturla H. Eik-Nes as my main supervisor was a bit scary. Would I ever be able to live up to his expectations? Sturla, your great professional capacity combined with your enthusiasm, humor and your claimed ‘telepathic, empathetic talents’ have certainly been both stimulating and reassuring. Your patience has indeed been much greater than mine, and your lack of willingness to compromise on quality has challenged me and taught me much about research and academic work. I will always be grateful for what you have done for me throughout these years.

My two co-supervisors, Professor Håkon K. Gjessing and Professor Per Grøttum have offered their vast knowledge of mathematics and informatics — always inspiring and educational. Håkon, you have ‘stuffed’ my Mailbox with tables and figures, looking for new ways in which you could explain your ideas. Your self-irony and sense of humor have encouraged me through many long evenings with the computer. Per, you took our data files to pieces and put them together again. Your comments to my manuscripts were always comprehensive and right to the point, and always served with a supporting smile. Thanks to both of you for never leaving me feeling stupid after having asked all those silly questions.

Many thanks to my other co-authors for valuable help and input. Thank you, Torbjørn Eggebø, for encouraging my research activity and for always being prepared to discuss results and presentations. Tore G. Bjåstad, thank you for all your help, and for genuinely wishing to introduce me to the ‘mystical minds’ of the ultrasound machines.

Nancy Lea Eik-Nes — your work with all my manuscripts has been impressive and very stimulating. I have always been looking forward to reading your comments and I have learned a lot from discussing language and scientific writing with you. Thanks for always being willing to help, and for your warm hospitality.

At Stavanger University Hospital, there are many people to thank. The former head of the department, Kjell Løvslett, initiated several Ph.D. projects before he retired, and this thesis would have been unthinkable without his efforts. I also want to thank Stein Tore Nilsen for always and without hesitating being prepared to help and find good solutions. Completing

this work would have been impossible without your support. At the Department of Obstetrics and Gynecology, I have appreciated the continuous backing from Sissel Moe Lichtenberg, Astrid Rygh, my doctor colleagues and the midwives. Leif K. Gjessing established the databases and supplied me with the data file for Study 2. Anita Steinbakk and Ane Cecilie Munk: it has been a pleasure and an inspiration to follow the progression of your Ph.D. studies. Thank you for valuable support and important discussions. To all my colleagues in the House of Research, who have been cheering me on: Of course, I very much appreciate your discreet and well intentioned questions about my irregular working hours….

The ultrasound measurements that have been analyzed in the 4 studies, were collected by ultrasound midwives in Trondheim, Stavanger and in Oppland County. Their understanding of the need for thorough data registration is exemplary, and the studies would have been worthless without the high quality of the data provided by these sonographers.

On my visits to Trondheim, the warm, welcoming atmosphere at the National Center for Fetal Medicine has always impressed me, and I am most grateful to *all* the staff for their positive attitude and their interest in my research. I have consistently got help, answers and good advice whenever I asked Gerd Inger Lånke, Eva Tegnander, Harm-Gerd Blaas and Anne Brantberg, or anyone else at the center. Morten Dreier: Thank you for always reducing the level of panic; your almost irritating calmness when data programs don't work and computers crash, has been most reassuring.

Warm thanks to the library at Stavanger University Hospital, Hilde Magnussen and Mara Andrejic, for finding all those 'impossible' articles so quickly.

To all my friends inside and outside the hospital: Thank you for encouraging me with phone calls or by dragging me out for coffee breaks, walks, wine or exercise when I did not think that I had time for it. I have needed and very much appreciated your friendship.

It is impossible to express what the support and love from my three fantastic daughters have meant to me throughout these last years. Ingrid, Kjersti and Turid: No thesis can compare to my maternal pride of you, and a doctoral degree would have been neither worthwhile nor possible without your kind help and understanding. Thank you for sharing with me so much of your joy, talents, time, company, and wonderful music!

List of papers

The thesis is based on the following papers:

I

Økland I, Gjessing HK, Grøttum P, Eik-Nes SH. Biases of traditional term prediction models: results from different sample-based models evaluated on 41 343 ultrasound examinations. Ultrasound Obstet Gynecol 2010; 36: 728–734.

II

Økland I, Gjessing HK, Grøttum P, Eggebø TM, Eik-Nes SH. A new population-based term prediction model vs. two traditional sample-based models: validation on 9046 ultrasound examinations. Ultrasound Obstet Gynecol 2011; 37: 207–213.

III

Økland I, Nakling J, Gjessing HK, Grøttum P, Eik-Nes SH. Advantages of the population-based approach to pregnancy dating demonstrated with results from 23,020 ultrasound examinations. Ultrasound Obstet Gynecol 2011. Accepted manuscript online: 25 Aug 2011. DOI: 10.1002/uog.10081.

IV

Økland I, Bjåstad TG, Johansen TF, Gjessing HK, Grøttum P, Eik-Nes SH. Narrowed beam width in newer ultrasound machines shortens measurements in the lateral direction: fetal measurement charts may be obsolete. Ultrasound Obstet Gynecol 2011; 38: 82–87.

Abbreviations

AC	Abdominal Circumference
AGA	Average for Gestational Age
BMUS	British Medical Ultrasound Society
BPD	Biparietal Diameter
CI	Confidence Interval
CRL	Crown-Rump Length
EDD	Estimated Date of Delivery
FL	Femur Length
GA	Gestational Age
HC	Head Circumference
ISUOG	International Society of Ultrasound in Obstetrics and Gynecology
IUGR	Intrauterine Growth Restriction
IVF	In Vitro Fertilization
LGA	Large for Gestational Age
LMP	Last Menstrual Period
MAD	Mean Abdominal Diameter
m/s	meter/second
NCFM	National Center for Fetal Medicine
OFD	Occipito-Frontal Diameter
PSF	Point-Spread Function
SGA	Small for Gestational Age

The myths and mysteries surrounding pregnant women and their childbirths have always puzzled the minds of investigators. 1900 years before the introduction of obstetric ultrasound, a Chinese mathematician, Sun-Tsu, probably in the first century AD wrote in a mathematics textbook:

A pregnant woman, who is 29 years of age, is expected to give birth to a child in the 9th month of the year. Which shall be her child, a son or a daughter? (Beckmann 1971)

And the solution, according to the Czech statistician and physicist Petr Beckmann (1924–1993), illustrates what he describes as a 'quaint mixture of mathematics and mumbo-jumbo' — giving rather little credit to the ancient philosopher:

Take 49; add the month of her child-bearing; subtract her age. From what remains, subtract the heaven 1, subtract the earth 2, subtract the man 3, subtract the four seasons 4, subtract the five elements 5, subtract the six laws 6, subtract the seven stars 7, subtract the eight winds 8, subtract the nine provinces 9. If the remainder be odd, the child shall be a son; and if even, a daughter (Beckmann 1971).

This takes Beckmann to the conclusion that 'to say very impressively nothing at all is the secret of all such oracles' (Beckmann 1971).

The actual 'mumbo-jumbo' was about the sex of the child. However, in all cultures at all times there has been an abundance of mysticism and old wives' tales speculating about the expected time of the birth of a child — in addition to assorted interpretations of signs indicating the sex, weight, talents, future health and predestined fate of babies about to be born. These aspects surrounding the birth of a child are still keeping "old wives" occupied, but are also puzzling the minds of excellent researchers in various disciplines. Some of these aspects have become the primary objectives of today's basic pregnancy care and fetal medicine.

Introduction

Antecedent theories on pregnancy length and time of delivery

Aristotle (384–322 BC)

The ancient Greek philosopher (Figure 1), considered a pioneer of the field of logic, formulated theories on the duration of gestation in different animals; he reasoned that there was a specific term that varied between animal species according to the size and life expectancy of the animal (Harvey 1847). However, he wrote, 'the human fetus is expelled both in the 7th and 10th month, and at any period of pregnancy between these; moreover, when the birth takes place in the 8th month it is possible for the infant to live' (Harvey 1847).

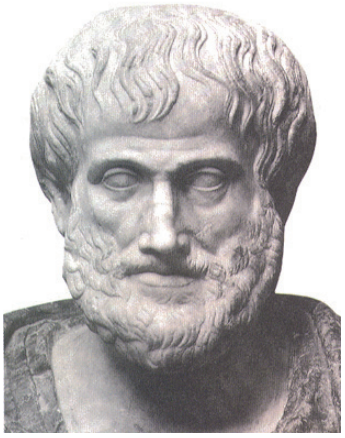


Figure 1. *Aristotle*

http://www.liberal-vision.org/2010/01/07/aristotle-384-322-bce-the-politics/ *(Accessed 25 May 2011)*

William Harvey (1578–1657)

William Harvey (Figure 2), the 'father of modern physiology', was an English physician and the first to describe human blood circulation in detail; and he also founded the new science of embryology through his studies on eggs. He studied Aristotle's hypotheses thoroughly and discussed them in his works. Harvey wrote that 'women are most prone to conceive either just before or just after the menstrual flux, for at these periods there is a greater degree of heat and moisture' (Harvey 1847).

Figure 2. *William Harvey*

http://en.wikipedia.org/wiki/File:William_Harvey_(_1578-1657).jpg#file *(Accessed 25 May 2011)*

When it came to the duration of human pregnancy, he rather cited the Bible (the nine months from the Feast of the Annunciation in March until Christmas Day (Luke 1: 26–38)).

However, he stated that births may take place before the 7th or after the 14th month — as 'the best ascertained signs of pregnancy have sometimes deceived not only ignorant women, but experienced midwives and even skilful and accurate physicians' (Harvey 1847).

Hermanni Boerhaave (1668–1738)

Boerhaave (Figure 3) was one of the most influential figures in the early modern period of medicine, and he has later been recognized as the founder of clinical bedside teaching and modern academic hospitals. In a collection of his academic lectures in five volumes, edited and annotated by Albert Haller and published in 1744 (Figure 4) (Boerhaave 1744, Hutchon 1998, Baskett and Nagele 2000), there is a lecture *On Conception* where he says: 'Women for the most part are impregnated after the end of their period. Numerous experiments undertaken in France confirm

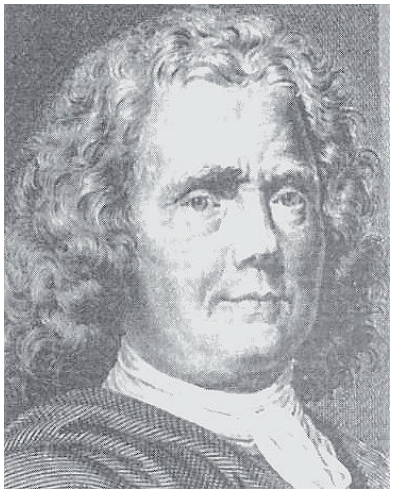


Figure 3. *Hermanni Boerhaave*

http://www.sciencemuseum.org.uk/broughttolife/people/hermanboerhaave.aspx *(Accessed 25 May 2011)*

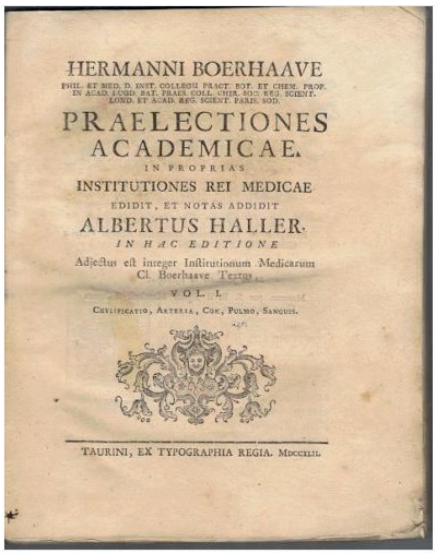


Figure 4. *Boerhaave's lectures*

http://www.baillement.com/lettres/boerhaave.html *(Accessed 25 May 2011)*

this; for of 100 births altogether, 99 came about in the 9th month after the last menstruation by counting 1 week after the last period and by reckoning the 9 months of gestation from that time. For, at that time the uterus is purged and empty, and the plethora are drained out' (Baskett and Nagele 2000).

There was no specification of whether 'counting 1 week after the last period' meant counting from the start or from the end of the menstruation. Nevertheless, this suggested rule of Boerhaave is the first known presentation of an algorithm for calculating the date of delivery.

Franz Carl Nägele (1778–1851)

Nägele, the professor in obstetrics at the University of Heidelberg (Figure 5) first published a method for calculating date of delivery 200 years ago (Nägele 1812), later known as Nägele's rule. When the rule is discussed today, Nägele's 'Lehrbuch der Geburtshülfe für Hebammen', which was printed in 14 editions after the first 1830-edition, is often cited (Nägele 1868). However, his original rule dates from 1812 (Figure 6) (Nägele 1812). He quoted Boerhaave's method, but added his own observations: 'According to experience, the woman in her reproductive phase does not always have the same capacity to conceive. The time the woman is most likely to conceive is immediately after menstruation'. He then cites Boerhaave in telling that women 'always conceive after the last menstruation and scarcely at any other time' (Nägele 1812, Baskett and Nagele 2000).

Figure 5. *Franz Carl Nägele*

http://www.123people.de/s/carl+nägele (Accessed 25 May 2011)

Nägele continued with his own observations: 'The usual calculation of the duration of pregnancy starting from

the last menstruation is correct in most instances; and conception within the last third of the cycle or in the second half between 2 periods is unusual and an exception to the rule' (Baskett and Nagele 2000).

Erfahrungen
und
Abhandlungen aus dem Gebiethe
der
Krankheiten
des
weiblichen Geschlechtes.

Nebst
Grundzügen einer Methodenlehre
der
Geburtshülfe
von
D. Franz Carl Nägele,
ordentlichem Professor der Arzneywissenschaft zu Heidelberg.

Mit vier Kupfertafeln.

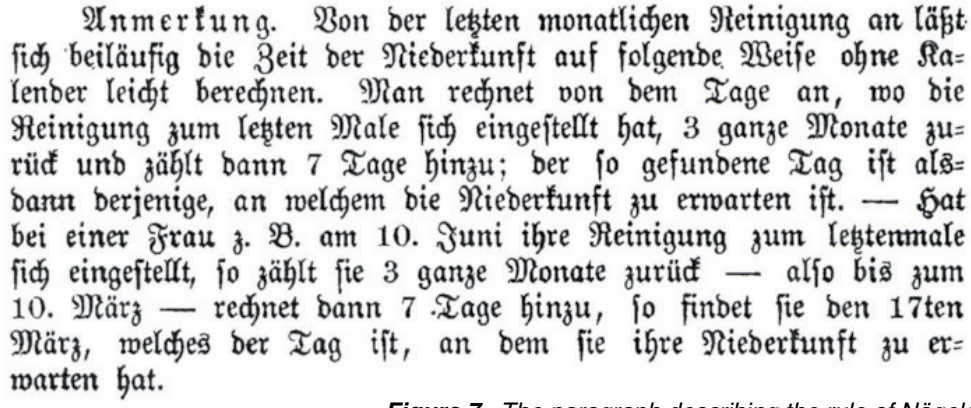
Mannheim,
bey Tobias Loeffler.
1812.

Figure 6. *The publication with Nägele's rule from 1812*

(Baskett and Nagele 2000)

In the 13th edition of his textbook for midwives (Nägele 1868), Nägele discussed the problems surrounding pregnancy dating, listing the different possibilities of estimating date of delivery by counting from 1) Conception 2) Absence of menstruation 3) Recognized fetal movements or 4) Increasing size of the uterus. Then he states that the time of conception remains hypothetical and that calculating from the last menstrual period (LMP) is the reasonable way of dating. In cases of unreliable LMP-data, the other methods of estimation have to be used, but then the day of delivery 'cannot be estimated precisely'. Afterwards, the exact way of calculating date of

delivery in the mode of Professor F. C. Nägele; counting 3 months backward from the LMP and then adding 7 days — the Nägele's rule — is accounted for (Nägele 1868):



Anmerkung. Von der letzten monatlichen Reinigung an läßt sich beiläufig die Zeit der Niederkunft auf folgende Weise ohne Kalender leicht berechnen. Man rechnet von dem Tage an, wo die Reinigung zum letzten Male sich eingestellt hat, 3 ganze Monate zurück und zählt dann 7 Tage hinzu; der so gefundene Tag ist alsdann derjenige, an welchem die Niederkunft zu erwarten ist. — Hat bei einer Frau z. B. am 10. Juni ihre Reinigung zum letztenmale sich eingestellt, so zählt sie 3 ganze Monate zurück — also bis zum 10. März — rechnet dann 7 Tage hinzu, so findet sie den 17ten März, welches der Tag ist, an dem sie ihre Niederkunft zu erwarten hat.

Figure 7. *The paragraph describing the rule of Nägele. (Nägele 1868)*

Hence, the algorithms of Boerhaave and Nägele differed in that while Boerhaave counted 9 months forward from 1 week after the last period, Nägele counted 3 months backward from the LMP and added 1 week. As we can see, Nägele was no more specific than Boerhaave on the issue of counting from the first or the last day of the LMP. Nevertheless, calculating from the first day has been the generally accepted method — obviously, it is easier to have a clear knowledge of the first, than of the last day of a period. However, in an amazing study from Belfast, Ireland (Gibson 1955), 6000 women were interviewed, and Gibson found that the common practice in Ireland was to calculate the estimated date of delivery (EDD) from the last, rather than from the first day of the LMP.

As will be discussed further on, many authors have claimed that the length of human pregnancy is longer than Nägele estimated, and the lack of accordance has led to speculations as to whether Nägele's rule was based on anecdotal evidence of the common belief that human pregnancy lasted for 10 menstrual cycles, rather than on empirical data (Mittendorf *et al.* 1990). Regardless of these reservations and whether the rule is to be attributed to Boerhaave (1744) or Nägele (1812), it served as a reference point for investigations into pregnancy length and EDD until the last decades of the 20th century.

Towards evidence-based medicine and modern technology

The first models for estimation of gestational age (GA) from fetal ultrasound measurements were introduced by Stuart Campbell (Campbell 1969). At that time, the ultrasound machines were developed from A-mode (one-dimensional amplitude imaging) to B-mode (two-dimensional images with brightness variation), producing images easier to understand and interpret. Still, the only meaningful fetal measurement to obtain was the biparietal diameter (BPD) (Figure 8).

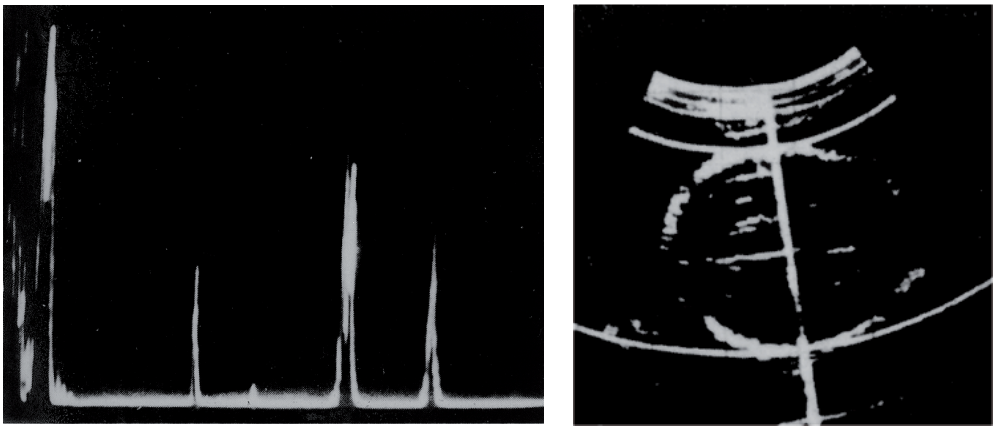


Figure 8. *To the left, a typical A-mode image showing the skull echoes and the smaller midline echo. To the right, a B-mode image of a fetal skull at the BPD measurement level.*

Campbell's BPD-measurements were achieved by combining A-mode and B-mode. He was able to determine the orientation of the fetal head and assure a correct transverse scan at the appropriate level by means of the 2 different scans being made at right angles to each other. This method increased the precision of the measurement technique, avoiding the problems resulting from head moulding and asynclitism (Campbell 1968).

Campbell's studies on fetal age and the measurement charts he developed have been fundamental in the development of modern fetal medicine. His description of the BPD measurement plane in order to obtain a correct ultrasonic measurement is still considered a reference article (Campbell and Thoms 1977), which is widely cited. Campbell prepared the way into the era of ultrasound-based pregnancy dating

through early fetal cephalometry, and introduced the idea of a routine fetal examination, with correct fetal age estimation and term prediction as the primary purposes (Campbell 1969). He then — somewhat reluctantly — found that the field of prenatal diagnosis was more closely connected to the dating examinations than he had assumed. Nevertheless, what he recently described as ‘an unexpected dilemma’ brought him into the fields of prenatal anomaly diagnosing with ultrasound (Campbell 2010), where he also became a true pioneer.

Norwegian ultrasound innovators

Campbell's first studies on fetal examinations were published around 1970, and at that time the ultrasound machines were, one by one, also being imported to Norway by committed obstetricians. The birth departments at the hospitals in Tromsø (Joachim Jenssen and Hans Andreas Sande) and in Bergen (Per Bergsjø and Christian Brodtkorb) were the first Norwegian departments to have their own scanners. These environments published smaller studies in Norwegian and Nordic journals in the early 1970s, concerning the use of and experience with obstetric ultrasound (Kvande 2008). Gradually, two different fetal growth curves were developed, by Bergsjø and Brodtkorb and by Sande (Kvande 2008). These curves were based mainly on longitudinal studies of third-trimester pregnancies, reflecting the fact that the primary aim of the curves was to diagnose intrauterine growth restriction (IUGR) in the third trimester (Bergsjø and Brodtkorb 1973).

According to the thesis by Lise Kvande (2008), there were 9 ultrasound scanners in Norway in 1974. In 1980, 40 departments had a scanner, and these 40 institutions took care of 83% of the deliveries. Internationally, the second-trimester routine examinations were about to be well established.

Sturla Eik-Nes started his work as a resident in obstetrics and gynecology in Ålesund in 1976 and a brand new ultrasound scanner arrived simultaneously.

Systematic measurements of second- and third-trimester fetuses were collected in Trondheim in 1979, and the basis for a new pregnancy dating wheel was formed, followed by a fetal growth formula. This was indeed a formula, based on polynomial regression analyses, not simply a collection of measurements from fetuses at different ages. The mathematical and medical skills of Per Grøttum were essential in transforming ultrasound measurements into a future prediction model for fetal age and growth (Eik-Nes *et al.* 1982a, Eik-Nes *et al.* 1982b, Eik-Nes and Grøttum 1983). Eik-Nes and Grøttum managed to provide Norway with ‘Snurra‘ (1983) — a pregnancy dating and fetal growth estimation model that was nationally used for at least 20 years.

During his years in Ålesund and the first years in Trondheim, Eik-Nes initiated two randomized controlled trials aimed at showing the potential benefits of routine second-trimester examinations (Bakketeig *et al.* 1984, Eik-Nes *et al.* 1984, Eik-Nes *et al.* 2000). A significant reduction of the need for induction in post-term pregnancies was a main finding. The introduction of ultrasound examinations to all pregnant women as a natural part of public pregnancy care was a hot topic at a scientific ultrasound meeting in Ålesund in 1984, and also at the Norwegian consensus conference on obstetric ultrasound, which was arranged 2 years later (Backe and Buhaug 1986, Kvande 2008).

Throughout more than a generation, Eik-Nes's continuous efforts to promote obstetric ultrasound in Norway as well as internationally, have been crucial for fetal medicine in general (Campbell 2010), for the organization of fetal routine ultrasound examinations in Norway, and for the education and training of midwives in obstetric ultrasound.

Pregnancy dating in Norway

During the decade before routine ultrasound examinations became a part of public health care in Norway, there was a complete lack of organization of the examinations and standardization of the dating models that were used. Whether an ultrasound-based EDD should be preferred to a reliable LMP-based date was an important part of the discussion. There were some 'local' curves in use at some hospitals, but the published curves were related to fetal growth in the third trimester and not to second-trimester dating. Hence, there was an urgent need for quelling the chaos and for increasing the impact of evidence-based medicine in Norwegian obstetric ultrasound.

The data that were collected from fetal examinations in Trondheim were used by Eik-Nes *et al.* to introduce the phrase 'TUL' — Term according to ULtrasound — which for 25 years has been a fixed Norwegian term, resulting in a uniform dating system from ultrasound, regardless of whether the LMP was reliable or not. Even the Medical Birth Registry of Norway, founded in 1967, from 1999 records the ultrasound-based EDD. In 1984, Eik-Nes and Grøttum developed an ultrasound form, called 'the blue form'. This was soon used nationally as a standard form for registration of the ultrasound findings throughout pregnancy, and it became part of the pregnant woman's medical journal until delivery.

Snurra ('Trondheim–1984')

The obstetric wheel 'Snurra' (Eik-Nes and Grøttum 1983), shown in Figure 9, was introduced in 1984 and was soon being used for ultrasound-based term prediction all over Norway. According to Snurra, the duration of Norwegian pregnancies was changed from 280 to 282 days; and somewhat surprisingly, without any objections from the obstetricians. EDD was predicted from BPD measurements between 38 and 60 mm, and fetal growth and weight could be estimated from BPD and mean abdominal diameter (MAD) measurements throughout the third trimester. The model was based on the fetal measurements from the Trondheim study; a population of 90

pregnant women with anticipated normal pregnancies, all of them carefully selected regarding menstrual history. The women were included in a prospective, longitudinal study, and measurements were taken from each fetus approximately 10 times. Fourth-order polynomial regression analysis was used to establish the curves.

Figure 9. *The pregnancy wheel 'Snurra'*

In Katarina Tunón's thesis from 1999, 5 studies on different aspects of ultrasound-based pregnancy dating were included (Tunón *et al.* 1996, Tunón *et al.* 1998, Tunón *et al.* 1999b, Tunón *et al.* 1999a, Tunón *et al.* 2000), and Snurra was the dating method that was used as the ultrasound-based term prediction model in all the studies. Tunón *et al.* were able to prove that ultrasound dating should be the method of choice even when LMP-data were reliable (Tunón *et al.* 1996), and in pregnancies conceived after in vitro fertilization (Tunón *et al.* 2000). They also stated that there was a tendency towards less precise predictions if the ultrasound examination was carried out earlier than pregnancy week 17–18; the model then estimated the date of delivery too late (Tunón *et al.* 1998). The authors' assumption was that the term prediction model needed to be improved. To anticipate the course of events; this came about 9 years later (Gjessing *et al.* 2007). However, a general recommendation emerged that the second-trimester ultrasound examination in general should not take place until after week 18. Moreover, this recommendation has been so faithfully followed that most dating examinations in Norway for the last 10 years have been performed closer to 19 weeks' gestation, rather than earlier. The women who proved to have received a 'too early' scanning appointment have, as a

rule, had an offer to come back a couple of weeks later. Overall, this routine has resulted in quite reliable EDD estimations from Snurra in large populations.

The Snurra model was one of the term prediction models evaluated in the studies that form the basis of this thesis — and thus the name 'Trondheim–1984' has been used. After having been the only ultrasound-based model used to date Norwegian pregnancies for 20 years, and one of two methods for almost 5 additional years, it has gradually been phased out of use during recent years, as a consequence of the introduction of a new population-based prediction model (Gjessing *et al.* 2007). The obvious advantages of 20 years of uniform dating across a country cannot be overemphasized when it comes to evaluating and comparing all kinds of perinatal outcome (Zeitlin *et al.* 2007, Salomon *et al.* 2005).

Even though Snurra has been shown to have certain shortcomings, the fact that it was constructed with the sampling and statistical methods of the 1980s, but still able to keep its position through more than two decades of continuous progress in all technological fields, should indicate that its high quality was not just a coincidence.

Terminhjulet ('Bergen–2004')

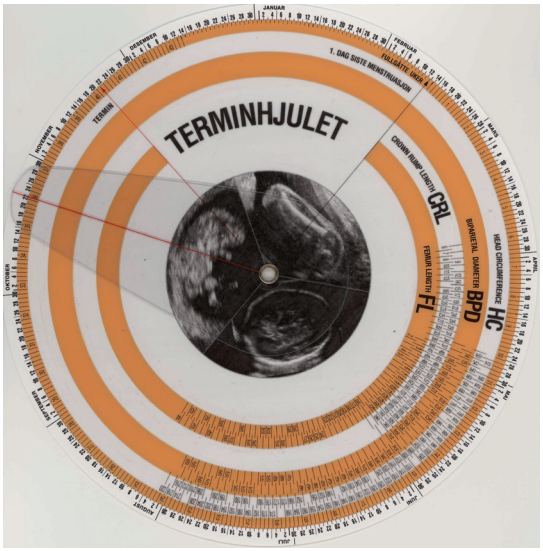


Figure 10. *The dating wheel 'Terminhjulet'*

A new dating method, shown in Figure 10, was introduced in 2004. The occasion was in part the results of a study by Kiserud and Rasmussen (1999), that confirmed the findings by Tunón *et al.* (1998), that Snurra underestimated fetal age with BPD measurements before pregnancy week 17–18 (Figure 11). The studies by Johnsen and coworkers made up the basis for this new obstetric wheel (Johnsen *et al.*

2004, Johnsen *et al.* 2005), which in our studies was given the more international name ‘Bergen–2004’. In the same way as Snurra, this model is also based on the statistical methodology for traditional dating models described by Altman and Chitty (1993, Altman and Chitty 1997, 1994) and Royston and Wright (1998).

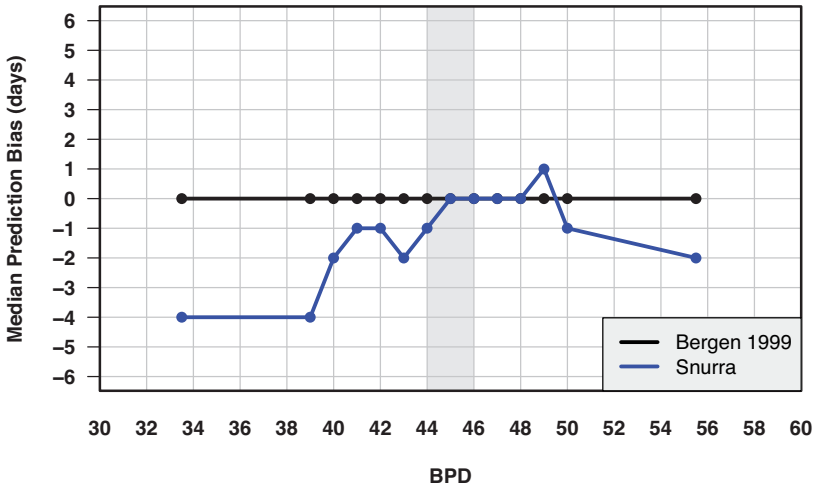


Figure 11. *Results from the evaluation of Snurra on a population of 8029 pregnancies in Bergen, shown with a figure by H.K. Gjessing.*

(Kiserud and Rasmussen 1999)

Fractional polynomial regression analysis was used to establish the prediction curves that provided the basis for Terminhjulet. EDD can be predicted from BPD measurements 14–60 mm, femur length (FL) 2–44 mm or head circumference (HC) 50–134 mm. The model was constructed from a prospective, cross-sectional study of 650 healthy women with regular menstrual periods and singleton, uncomplicated

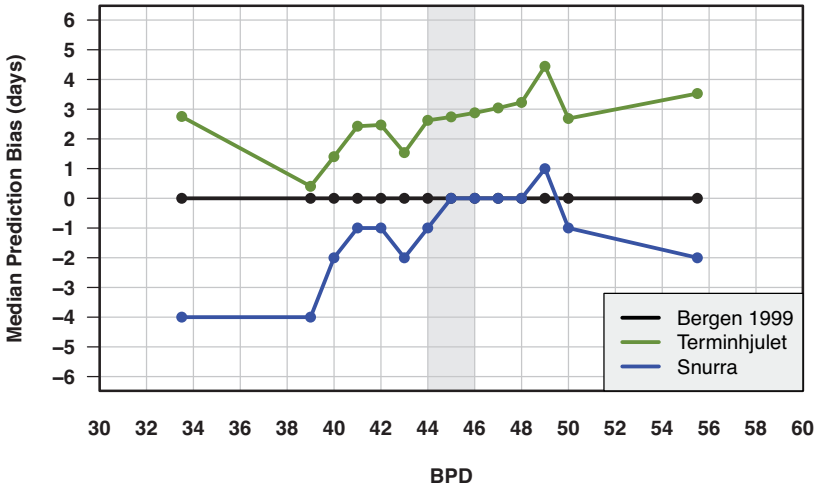


Figure 12. *Results from the evaluation of Snurra on a population of 8029 pregnancies in Bergen, shown with a figure and the additional data for Terminhjulet by H.K. Gjessing.*

(Kiserud and Rasmussen 1999)

pregnancies. The study population was included and examined at Haukeland University Hospital in Bergen, Norway, during the years 2001–2003.

Bergen–2004 was gradually adopted by some Norwegian hospitals in the years after 2005. The impression was given that this new model would perform far better than the older model (Snurra), regarding both term prediction and fetal growth assessment. However, there were many skeptics who wondered whether Terminhjulet, indeed with a new data sample but with rather 'old-fashioned' statistical methods, could be the solution to Snurra's limitations.

An evaluation of the Bergen–2004 model comparing it with Trondheim–1984 was carried out on a population of 11 238 pregnancies from the Oppland County in Norway (Backe and Nakling 2006). This evaluation confirmed the underestimation of fetal age if Trondheim–1984 was used in early second-trimester predictions and they concluded that the new reference values more precisely assessed GA. However, there were some shortcomings in the study; these concerned the inclusion of women with reliable LMP data only, the use of the mean instead of the median bias as an outcome measure and the complete lack of acknowledgment of the fact that Bergen–2004 actually overestimated the pregnancy length and predicted the EDD too early (Eik-Nes *et al.* 2006), as shown in Figure 12.

Looking back, it is difficult to understand how the systematic prediction error of Bergen–2004 could appear undetected. In the study by two of the authors (Kiserud and Rasmussen 1999), where they evaluated Trondheim–1984 on 8029 ultrasound examinations, they also applied the prediction model by Altman and Chitty (1997) to BPD measurements from their own study population, for comparative purposes. They then concluded that the Altman and Chitty model was not a suitable dating model for Norwegian pregnancies, because of — as they emphasized — the systematic overestimation of pregnancy length and the prediction of EDD 3–4 days too early. However, 5 years later, when the new Bergen–2004 dating model was published (Johnsen *et al.* 2004), they compared their brand new dating curves with the one by

Altman and Chitty (1997), and apparently considered it an advantage that the curves were almost overlapping — with differences in EDD predictions of only 1 day or less. Hence, the predictions from the new Bergen–2004 model are more or less identical to the predictions from a model that some years earlier had been considered not suitable in Norway, due to erroneous predictions.

The prediction model that was meant to answer the question in the title of the 1999-study (Kiserud and Rasmussen 1999): 'Assessment of gestational age using ultrasound — can the method be improved?' turned out, unfortunately, not to be a convincing improvement (Figure 12).

eSnurra ('Trondheim–2007')

The study that is the basis for the new population-based term prediction model, eSnurra (Figure 13), was published in 2007 (Gjessing *et al.* 2007). This was a new cooperation project between Eik-Nes and Grøttum, now in alliance with the mathematician Håkon K. Gjessing. To base a term prediction model on fetal ultrasound measurements from a non-selected population of 36 982 pregnancies (41 343 ultrasound examinations), was a new approach that seemed obvious to statisticians and mathematicians, but not equally obvious to obstetricians. The intrinsic difference is to use ultrasound measurements from a given point of time during the second trimester, and to combine the data from these with the actual time of the delivery. This makes it possible to calculate the median remaining time of pregnancy and the EDD. Conversely, the traditional, sample-based models use the fetal measurements to estimate the LMP — the EDD is then found by adding a specified number of days to the estimated LMP date. This contrast has been rather complicated to communicate, even though it is both evident and logical (Hutchon and Ahmed 2001, Salomon *et al.* 2010, Gjessing *et al.* 2007).

The huge amount of data being collected and processed with a population-based model, opens for the application of more advanced and robust statistics than was possible with the polynomial regression models that were used to develop the

traditional models (Altman and Chitty 1994). The obvious statistical approach to the population-based estimations is to use a non-linear quantile regression model (Gjessing *et al.* 2010), of which many are available. The Norwegian model used a local linear quantile regression (LLQR) model (Gjessing *et al.* 2007), while a recent, corresponding French model used another variant, namely a spline-smoothed quantile regression (Salomon *et al.* 2010).

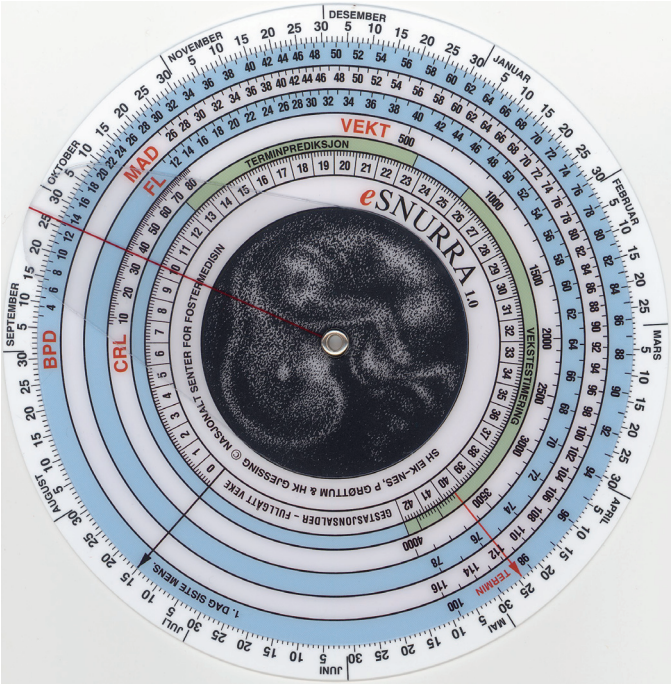


Figure 13. *'eSnurra'*

eSnurra predicts the EDD from BPD 25–60 mm or from FL 11–42 mm. The original data included fetal measurements from ultrasound examinations of unselected, singleton pregnant women. Pregnancies complicated by stillbirth, anomalies, possibly

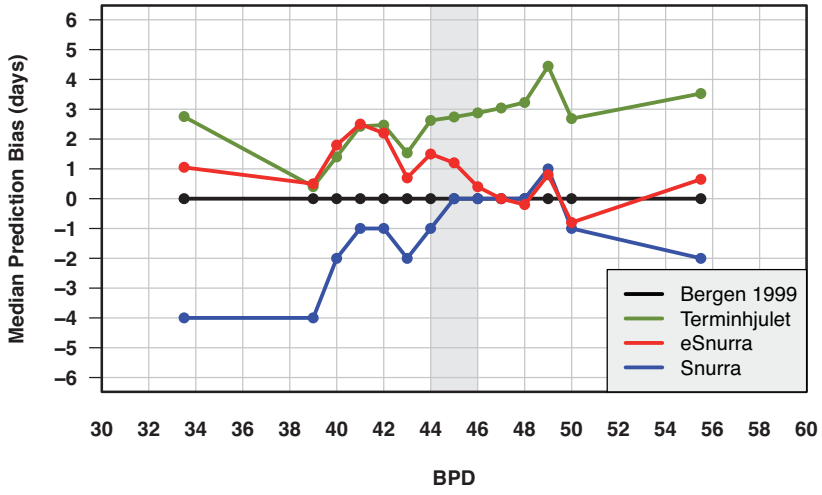


Figure 14. *Results from the evaluation of Snurra on a population of 8029 pregnancies in Bergen, shown with a figure and the additional data for Terminhjulet and eSnurra.*

(Kiserud and Rasmussen 1999)

abnormal fetal growth or induction of labor for reasons other than post-term pregnancy, were excluded (Gjessing *et al.* 2007).

Traditional sample-based models that are based on limited sample sizes have to be evaluated on population data to assess prediction quality. With a population-based model, the performance of the method is evaluated directly from the reference material. However, if the aim is to get a measure of how the model's predictions suit other populations with possibly different examination routines, a validation on the actual population has to be carried out. Figure 14 shows how eSnurra suits the study population from Bergen.

From Nägele to ultrasound

The rule not as valid as it seemed to be

First, every doctor and midwife who has tried to determine modern women's LMP, knows how difficult it is to achieve so-called 'reliable' data (Geirsson and Busby-Earle 1991). In his study, Gibson (1955) had to exclude 1/6 of the 6000 interview cases by reasons such as: (1) A fabrication of the 'dates' to agree with a marriage *(sic)* early in pregnancy. (2) Uncertain estimate for different reasons, among them 'lack of ordinary intelligence *(sic)* on the part of the patient'. We do not know how Nägele and Boerhaave solved this 2–300 years ago. Nevertheless, their algorithm was helpful — even with several inherent assumptions — so obviously the rule represented the method of choice for a very long time.

Regular menstrual cycle: Nägele's rule is based on ovulation and fertilization occurring on cycle day 14, and on a cycle length of 28 days. If ovulation occurs earlier or later or the periods are irregular, the rule becomes inaccurate.

Varying length of the follicular phase is not uncommon (Baird *et al.* 1997), neither are menstruation-like bleedings in early pregnancy, which can be mistaken for a period(Gjessing *et al.* 1999). Another confounding factor — at least for some women — is the varying time between coitus and the fertilization (Perloff and Steinberger 1964, Wilcox *et al.* 1995, Robinson 1973).

Length of calendar months: If a rule like this should operate precisely, one must assume an equal number of days in all twelve months amounting to 30.42 days each (365 days/12), resulting in an average pregnancy length of 280.78 days according to Nägele ((30.42 days x 9) + 7). The fact that the months actually are of unequal length brings about a variation in the calculated pregnancy duration, depending on the month of the year the LMP takes place. For example, the LMPs in the first 3 weeks of May will include 6 'long' months of 31 days; May, July, August, October, December and January, and will simultaneously avoid the short February month. A pregnancy with due date in the middle of February is therefore assumed to last for 283 days.

The ‘shortest’ pregnancies arise in years that are not leap years, for instance with LMPs in June or September — the estimated pregnancy length will then be 280 days.

Studies indicating that a normal pregnancy probably lasts longer than 280 days from a reliable LMP (Mittendorf *et al.* 1990)(Bergsjø *et al.* 1990, Tunón *et al.* 1996, Gjessing *et al.* 2007) actualized speculations on whether Nägele really meant counting from the first or the last day of the LMP. Some authors have brought Nägele's rule back into the discussions on term prediction (Olsen and Aaroe Clausen 1997, Nguyen *et al.* 1999, Baskett and Nagele 2000). However, at the moment it seems more reasonable to refine the ultrasound dating models than to resuscitate Nägele's rule by adding 10 days instead of 7, or counting from the last day rather than from the first. It appears evident that the LMP should be used only in the scheduling of the dating scan (Gardosi *et al.* 1997, Bottomley and Bourne 2009).

Ultrasound dating takes over

Throughout most of the 1990s, there were strong and differing opinions among obstetricians on how to reliably date pregnancies. The routine ultrasound examinations were fairly well established in most western countries, but there was less agreement on what to do if the ultrasound-based EDD differed from the LMP-based date.

When Tunón et al. (1996) were able to show that ultrasound dating was superior to dating even from *reliable* LMPs in a large population, they indeed answered many questions on how to use the ultrasound dating methods. In Norway, this study supported the established policy of using the term according to ultrasound, ‘TUL’, independent of the LMP term date. Internationally, the more ‘conservative’ view still prevailed, claiming that the LMP-method was the most trustworthy (Rossavik and Fishburne 1989). Rossavik and Fishburne found no significant difference between the ultrasound-based and the LMP-based EDD. However, their findings were based on a comparison of only two very strictly selected small groups (28 women with IVF-

pregnancies vs. 60 with highly reliable LMP-data). Nevertheless, this article remained a reference against routine ultrasound dating for a long time, not least because they introduced the principle of not necessarily changing the date of the EDD; 'if a pregnant woman has fairly regular periods and knows her LMP within a time-frame of ± 1 week, dates should not be changed unless the discrepancy between menstrual age and ultrasound age is 14 days or more' (Rossavik and Fishburne 1989).

The policy of not changing the EDD if the discrepancy was less than 4, 7, 10 or 14 days was soon adopted and used for many years (Gardosi *et al.* 1997, Gardosi 1997, Blondel *et al.* 2002). This may be regarded as a kind of a compromise between dating with ultrasound or dating according to Nägele. In my opinion, similar to many other compromises it became a neither–nor solution, creating confusion and concern among the mothers-to-be, and making comparisons of perinatal outcome measures related to GA problematic.

Little by little, agreement that the ultrasound dating models were superior to LMP-dating in most situations emerged (Kieler *et al.* 1995, Mongelli *et al.* 1996, Mul *et al.* 1996, Gardosi and Geirsson 1998, Nguyen *et al.* 1999, Tunón *et al.* 1996). Thus, during the last decade the discussion has been about when, what and how to measure (Persson 1999, Taipale and Hiilesmaa 2001, Källén 2002, Saltvedt *et al.* 2004, Verburg *et al.* 2008b) — and which ultrasound-based dating model to use (Mongelli *et al.* 2003, Sladkevicius *et al.* 2005, Salomon *et al.* 2010). Moreover, at least in Europe, it became common practice to use the EDD determined at the ultrasound examination, no matter how 'certain' the reported LMP might be (Taipale and Hiilesmaa 2001, Bottomley and Bourne 2009, Verburg *et al.* 2008b).

The ultrasound method

Immediate consequences

Ultrasound dating of pregnancies reduces the number of inductions of labor due to post-term pregnancy (Whitworth *et al.* 2010). This was originally shown in the Norwegian randomized controlled study addressing the routine use of ultrasound (Eik-Nes *et al.* 1984, Eik-Nes *et al.* 2000). The sample size of this study was based on an anticipated 50% reduction of inductions of labour in post-term pregnancies, resulting from the more reliable ultrasound dating. Reduced post-term rates, by up to 70%, have later been repeated findings in studies comparing ultrasound and LMP-dating (Persson and Kullander 1983)(Geirsson 1991, Tunón *et al.* 1996, Mongelli *et al.* 1996, Gardosi *et al.* 1997, Savitz *et al.* 2002, Blondel *et al.* 2002, Yang *et al.* 2002, Zeitlin *et al.* 2007).

The primary reason for the reduced post-term rate in pregnancies dated with ultrasound, is the varying interval between LMP and ovulation; the follicular phase (Baird *et al.* 1997). A delayed ovulation is much more common than an early one, therefore, the birth distribution curves with LMP-based EDDs are heavily skewed to the right and towards longer gestations when compared with the ultrasound EDD-based curves (Tunón *et al.* 1996, Yang *et al.* 2002, Savitz *et al.* 2002). Thus, after an ultrasound dating examination, the EDD is more often set to a later date than an earlier one (Gardosi *et al.* 1997); 60% vs. 33%, according to Tunón *et al.* (1996). This is also shown in a study comparing LMP and basal body temperature-rise as starting points for onset of gestation (Boyce *et al.* 1976); most of the variation in those delivering post-term could be ascribed to the varying LMP-to-ovulation interval.

Yang and coworkers (2002) stated that the overall overestimation of pregnancy length according to a reliable LMP was in the range of 3.3 days of the ultrasound-based estimate, and that this was a systematic, not a random error. In addition to showing the well-known decrease in post-term rates, they also studied the impact of ultrasound dating on the preterm birth rates. Because there are more downward

reclassifications of GA than there are upward, and because deliveries towards the pregnancy weeks 36–37 are frequent (the birth distribution curve is steeper towards term), dating with ultrasound not only reduces post-term rates, but also significantly increases the preterm ones (Goldenberg *et al.* 1989, Høgberg and Larsson 1997). These reclassifications have important public health and clinical implications (Yang *et al.* 2002), even though this fact is not being acknowledged among clinicians or researchers in perinatal epidemiology.

Challenges to ultrasound-based dating

Traditionally, the dating examination has been performed in the second trimester, and one single routine examination around week 18 is still what the Norwegian public health system offers pregnant women. Originally, the purpose of first-trimester screening examinations was not primarily the assessment of GA (Snijders *et al.* 1998), but later on, several studies indicating benefits from first- rather than second-trimester dating have been published (Bennett *et al.* 2004, Caughey *et al.* 2008) (Saltvedt *et al.* 2004, Verburg *et al.* 2008b, Bottomley and Bourne 2009). First-trimester dating recommendations are now also included in the NICE guidelines (2010). Probably, the potentially more accurate early dating is associated with the uniform fetal growth pattern throughout the first trimester, which is independent of ethnicity and other maternal/fetal factors (Pedersen *et al.* 2008b, Drooger *et al.* 2005). However, it is not equally obvious that more precise estimation of the time of conception and of fetal age necessarily implies a more accurate prediction of date of delivery (Gjessing *et al.* 2010).

Inverting the ‘Pyramid of prenatal care’ as suggested by Kypros H. Nicolaides (Nicolaides 2011c, Nicolaides 2011b), precisely illustrates the progress of ultrasound both in predicting and defining the risks of a variety of adverse pregnancy outcome, as well as in assessing the GA already in the first trimester. An integrated first hospital visit at 11–13 weeks, where the ultrasound findings are combined with data from maternal history and characteristics, as well as with findings of biophysical and

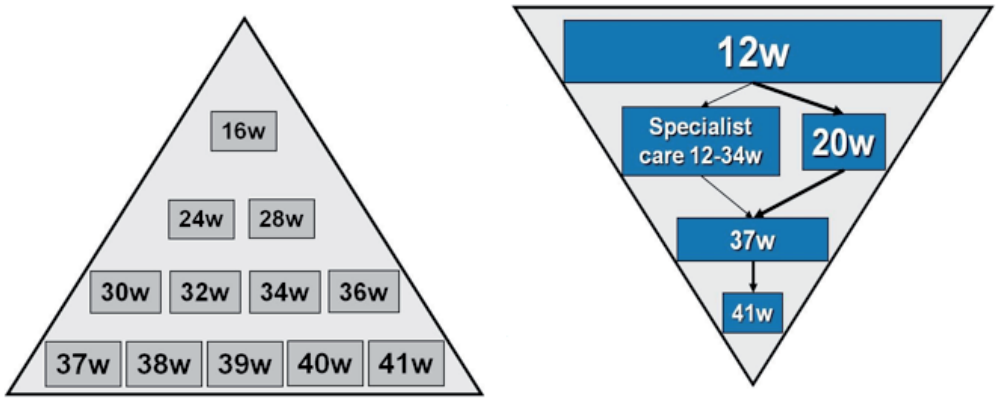


Figure 15. *The pyramid of traditional prenatal care (left) versus the proposed new pyramid (right) shifting the emphasis of pregnancy care and visits to the first trimester.*

(With kind permission from Prof Kypros Nicolaides and the Fetal Medicine Foundation (2010a))

biochemical tests, will move the emphasis of modern pregnancy care from the third to the first trimester, as shown in Figure 15.

The importance of an accurate EDD for the great majority of women who deliver within ±1 week of the estimated due date, is probably somewhat exaggerated in the minds of these mothers to-be. However, if complications arise at any point of time outside the around-term weeks, an exact knowledge of the correct GA/EDD may be essential for making the correct clinical decisions (Pexsters *et al.* 2010, Dias *et al.* 2011). 40 years after the first studies on ultrasound-based fetal age assessment, it remains a reality that 'Accurate gestational dating is one of the most important assessments obstetrical providers make in pregnancy, given that all of the various management strategies are dependent on knowing where the patient is in gestation' (Gottlieb and Galan 2008).

Apart from the significant reduction in post-term inductions of labor, pregnancy dating with ultrasound has additional advantages (Dias *et al.* 2011). The first-trimester screening tests, as well as the scheduling of routine examinations, invasive procedures and potential interventions up to a GA where the fetus is assumed to be capable of survival, have traditionally been based on GA estimates calculated from the LMP. After a gestation of 23–26 weeks the focus gradually becomes more

concentrated on the EDD, as typically demonstrated in discussions on post-term pregnancies, where the issue is how many days past the predicted EDD is the appropriate and acceptable time for induction of labor. Consequently, the terms 'assessment of GA' and 'estimation of date of delivery' refer to totally different times in the course of a pregnancy — although they often appear to be considered synonyms. The contrasting aims of the sample-based term prediction models (calculate GA from a hypothetical LMP/conception) and the population-based models (direct prediction of EDD) is thus important to bear in mind in discussions of GA and EDD. With a traditional model, knowledge of the exact pregnancy length is needed to calculate the EDD (LMP + pregnancy length = EDD), while with a population-based model, knowledge of the exact pregnancy length is necessary in order to compute the actual GA (pregnancy length – remaining time of pregnancy = GA).

If the LMP date is used only for scheduling of the dating examination, all questions where the length of gestation influences the answer will relate to the ultrasound-based GA/EDD after the examination is performed. Of course, the selection of a pregnancy dating model has an impact on the management of a pregnancy if complications arise in the weeks and months after the examination. The prenatal diagnostics and first-trimester examinations primarily relate to the GA, while the EDD concerns the follow-up of all kinds of complications in the second half of the pregnancy. For the extreme pre-term and post-term fetuses, a displacement of the EDD in the range of 2 days may have implications for both management and survival, in addition to the obvious effect on public health outcome measures (Lynch and Zhang 2007, Zeitlin *et al.* 2007).

Preterm deliveries: Preterm deliveries between completed pregnancy weeks 22 and 27 are associated with an almost day-by-day reduction in perinatal mortality (Markestad *et al.* 2005). In a recent Swedish study (Fellman *et al.* 2009), the overall perinatal mortality in the study cohort is 45%; it was reduced from 93% at 22 weeks to 66% at 23 weeks, and further lowered to 24% at 26 completed weeks. Obviously, such figures are basic components of our information to the parents and may act as

a strong motivation to maintain optimism and patience through very critical days. Moreover, the management schemes and clinical guidelines concerning threatening preterm deliveries (*e.g.* choice of tocolytic treatment, when to give antenatal corticosteroids, whether transfer to another hospital), delivery method (vaginal or Cesarean section) (Grant *et al.* 1996), and the neonatal management (decisions concerning the level of and the establishing and discontinuing of pediatric actions and operations), are all influenced by the knowledge of an exact GA (Miljeteig *et al.* 2007, Saugstad 2005, Pignotti 2008). Consequently, a difference of one or two days around pregnancy week 23–24 may alter the activity level surrounding the extremely preterm delivery, and hence the prognosis for the preterm born infant, crucially.

Post-term pregnancy: While the preterm deliveries are mainly unavoidable even if occasionally scheduled, iatrogenic post-maturity may follow an erroneously predicted term date. In the same manner as guidelines for treatment of extremely preterm born infants assume a precise estimate of fetal GA, the post-term induction recommendations should have their basis in reliable predictions of due date (Gülmezoglu *et al.* 2009, Mandruzzato *et al.* 2010). Because the EDD is model-dependent, recommendations on when to induce in post-term pregnancy is futile without a uniform system for pregnancy dating (Blondel *et al.* 2002, Zeitlin *et al.* 2007, Lynch and Zhang 2007). Different definitions of standard pregnancy length and term prediction models that vary in predictive quality may add up to variations in EDDs of several days, which of course influences both the estimations of risk associated with post-term pregnancy and the numbers of women needed to treat, *i.e.,* to induce.

While the World Health Organisation's definition of post-term pregnancy remains at 294 days, most western countries that have evidence-based guidelines concerning this issue, now recommend induction at 41 completed weeks, *i.e.,* 287 days (Gülmezoglu *et al.* 2009, Norwitz 2011). The NICE guidelines (2010) also state that ‘women with uncomplicated pregnancies should be offered induction of labor beyond 41 weeks’.

The incidence of post-term deliveries (≥294 days of gestation) varies in different studies, but with ultrasound dating it is generally below 4% (Hilder *et al.* 1998, Taipale and Hiilesmaa 2001, Wennerholm *et al.* 2009). Biased predictions of due date or adoption of a policy of earlier post-term induction may shift the deliveries from pregnancy week 42 to week 41, implying a 2–4 fold increase in induction rates, depending on the chosen cut-off. In addition to making comparisons of important perinatal quality indicators impossible (Zeitlin *et al.* 2007), 'unnecessary' inductions of labor in false positive post-term pregnant women put an extra strain on delivery units (Wennerholm *et al.* 2009), independent of the discussions on when to induce in post-term pregnancy. Smaller randomized studies have not shown increased risks related to induction of labor in post-date pregnancies (Heimstad *et al.* 2007), however, large register-based studies reveal several risk factors associated with post-term inductions in, for example, primiparous women or in women with advanced age, unfavorable cervix or a high body mass index (Rossen *et al.* 2010, Roos *et al.* 2010). Consequently, an obvious condition in discussions on induction routines is unbiased and uniform EDD-predictions with comparable post-term rates.

IUGR: Intrauterine growth restriction (IUGR) is another clinical pregnancy problem in which the diagnosis, the follow-up and finally the timing of the delivery, are based on a correctly predicted GA/EDD (Baschat 2004, Maršál 2009). The identification of a small for gestational age (SGA) fetus may follow from a significant discrepancy between a reliable LMP-based and the ultrasound-based EDD at the second-trimester examination (Tunón *et al.* 1999a, Nguyen *et al.* 2000, Källén 2004, Morin *et al.* 2005). However, the postponing of an ultrasound-based EDD relative to the LMP-predicted date is in most cases due to a late ovulation, and the majority of these pregnancies prove to be completely normal, with deliveries of normal infants close to the ultrasound-based due date (Tunón *et al.* 1999b, Larsen *et al.* 2000, Morin *et al.* 2005, Fox *et al.* 2008, Thorsell *et al.* 2008). First-trimester examinations have become an option in most European countries, and there are studies showing that even a first-trimester discordance between GAs calculated from LMP and ultrasound

may indicate an imminent growth restriction, or an even more severe adverse outcome (Bukowski *et al.* 2007)(Pedersen *et al.* 2008a, Pedersen *et al.* 2008b, Bottomley and Bourne 2009, Kirkegaard *et al.* 2011). As these examinations become more widespread, repeated studies will most likely confirm that an increased risk of IUGR/SGA, preterm delivery or preeclampsia may be identified with first-trimester examinations. This is consistent with the inverted pyramid in Figure 15.

Since traditional, sample-based term prediction models are based on estimating the LMP, it has been implied that these methods would provide EDDs and GAs that correspond more closely to those computed from LMP, than do the population-based models. Unfortunately, one single ultrasound examination cannot in itself identify the possibly early IUGR fetuses, irrespective of prediction model. A measurement, for instance a BPD of 45 mm, taken at the ultrasound examination is seen in relation to a dating chart and results in a prediction of a corresponding due date without paying attention to the LMP or other factors. Of course, an IUGR fetus sometimes has associated findings, like Doppler abnormalities, reduced amniotic fluid volume and asymmetric head/abdominal measurements. However, the conclusion of an IUGR diagnosis as opposed to an average-sized fetus of a younger age than expected, takes at least two examinations. An awareness of a significant difference between reliable LMP-based and ultrasound-based EDD-dates is nevertheless sensible, and in such cases further evaluations of fetal growth are essential in order to reduce the risk of erroneous dating of fetuses with a deviating growth pattern. All the same, the EDD should be changed according to the ultrasound measurements.

In addition to the benefits of the more accurate ultrasound dating, the ultrasound examinations in themselves have values that are considered essential and independent elements in modern pregnancy care; the anomaly screening with its panorama of results, dilemmas and treatment possibilities, the earlier diagnosis and follow-up of multiple pregnancies, the localization of the placenta, and the chances of collecting into specialist care mothers with risk factors indicating imminent maternal pregnancy complications, preterm labor, drug addiction or other social problems.

Drawbacks and disagreements

In a perfect world, a health system offering routine ultrasound examinations to whole populations of pregnant women should ensure that that the examinations met certain criteria related to standardization regarding the education of the sonographers, and the quality of ultrasound equipment and measurement charts. Regular assessments of quality should be carried out. Unfortunately, the real world is not perfect.

In Norway, the ultrasound examinations are for the most part performed at hospitals by specially trained midwives, who are educated at the National Center for Fetal Medicine in Trondheim. In addition, gynecologists and midwives in private practice and at hospitals carry out varying shares of the examinations, depending on geography and special interests.

There is no official system of certification for performers of the second-trimester scans (Dudley and Chapman 2002, Salomon *et al.* 2008, Ville 2008), although these measurements may influence outcome and in some cases have an impact on considerations leading to interventions. This is in contrast to the standardized routines and certification system surrounding the 11–13 weeks scan in the Fetal Medicine Foundation system (2010b, Nicolaides 2011a), where quality control studies in prenatal sonography have long been based on detection rates of fetal abnormalities (Snijders *et al.* 1998); there are no analogues to this within the field of pregnancy dating and fetal biometry (Salomon *et al.* 2006a).

Ultrasound societies such as ISUOG (International Society of Ultrasound in Obstetrics and Gynecology) (Salomon *et al.* 2011), and BMUS (British Medical Ultrasound Society) (Loughna *et al.* 2009) have published practice guidelines for assessment of fetal size and dating, where structures recommended to be measured and the technique describing how to measure them, are specified. Unfortunately, such guidelines have a tendency to be released many years too late to really be implemented in common practice, because fulfilling these new standards would involve a rejection of the already established, and apparently well-functioning local practices. It takes strong evidence to change clinicians' minds.

Fetal measurement charts

Strict selections versus population-based

Traditional reference charts for ultrasound dating of pregnancies are based on a calculation of GA from an estimated LMP, via a regression of gestational age (as a dependent variable) on the ultrasound measurements of fetal size (independent variable). The estimated LMP is derived by plotting a measurement value from a fetal examination on a standardized chart, which indicates the most likely fetal age-for-size at any given time. The EDD is then determined by adding the assumed pregnancy length (in days) to the estimated LMP date.

The normal range curves are based on menstrual data from selected women. Therefore, it is rightly claimed that these charts will never be more accurate than were the LMPs of the women in the study group, from whom the charts were computed (Campbell *et al.* 1985, Bergsjø *et al.* 1990, Hall 1990, Geirsson 1991, Hutchon 1998). (Interestingly, even inclusion criteria like 'intelligent and well motivated women' have been mentioned as selection criteria (Robinson 1973).) For most of the traditional models, selection was limited to only those women who were considered to have less uncertainty and variability related to their LMPs than what would be found in an unselected population. It is likely that their LMP data have a more narrow distribution around a 'true' mean, but nevertheless, the selection process in itself is a significant source of error.

In traditional charts for evaluation of fetal growth, it is of course the fetal size that is the dependent variable, being regressed on the GA as the independent variable.

The traditional approach to construction of reference charts results in a varying predictive quality (Mul *et al.* 1996, Tunón *et al.* 1998, Saltvedt *et al.* 2004), and the inherent weaknesses have been explained in various ways. Obviously, the selection of healthy, 'hyper-normal' women with reliable LMPs and normal pregnancies (Gjessing *et al.* 2007, Salomon *et al.* 2010) results in study groups that differ greatly from the populations of women who are being examined at the typical routine

examinations. The issue of representativeness will always be crucial in traditional, small prospective studies.

Another source of error is the assumed pregnancy length that is added to the estimated LMP. As a fixed number of days is used in every single pregnancy, a systematic bias will appear if this number is not correct for the population concerned (Nguyen *et al.* 1999, Gjessing *et al.* 2007, Salomon *et al.* 2010). In addition, the traditional models are often based on measurements performed by only one or two dedicated sonographers, and many authors consider this an important cause of the systematic bias observed with these models (Kurmanavicius *et al.* 1999, Salomon *et al.* 2006b, Gjessing *et al.* 2007).

The false impression has been that the fetuses, the mothers, as well as the examiners had to be hyper-selected if a reliable dating model was to be constructed; for many years, this led researchers to avoid the much greater and unselected population-based materials (Gjessing *et al.* 2007). Attempts to solve the problems concerning the 'never trust-worthy' LMP information were tried being solved through comparisons with IVF pregnancies — thereby establishing IVF pregnancies as the gold standard for the timing of events in normal pregnancies (Rossavik and Fishburne 1989, Mul *et al.* 1996, Mongelli *et al.* 2003, Mongelli *et al.* 2005). Reports in later years have caused concern, showing that assisted fertilization is associated with increased risks of adverse outcome (Jackson *et al.* 2004). Probably, these complications can be attributed to the factors leading to the infertility, rather than to the IVF in itself (Romundstad *et al.* 2008). Consequently, it has not been confirmed that the IVF pregnancies can reasonably be regarded as representative of normally conceived pregnancies when it comes to assessment of dating or growth (Chalouhi *et al.* 2011).

As the sample-based studies exclude women with unreliable LMP data, a self-contradictory consequence is that the considerable share of pregnant women for whom ultrasound dating is the only feasible alternative is excluded. With population-

based data and direct estimation of date of delivery, independent of GA and LMP data, such problems are avoided. In addition, it also provides an estimate of the uncertainty of the prediction (Salomon *et al.* 2010).

Another advantage of population-based term prediction models, is that inductions of labor by reason of post-term pregnancy do not have to be excluded from the analyses. As these inductions seem to take place at yet earlier GAs, a larger share of them are normal pregnancies, which are just not approved waiting for spontaneous labor to occur. Using the median, the term predictions are not influenced by the inductions taking place at 41 or 42 weeks — and the impact of the left-skewed birth distribution curve caused by the spontaneous preterm deliveries are also avoided (Gjessing *et al.* 2007, Salomon *et al.* 2010). Whether delivery comes about 2 days or 2 weeks before or after term does not affect the median, however, inductions of labor for all other reasons than post-term pregnancy, and all elective Cesarean sections must be excluded. The median effectively divides the population of births into two: half of them occurring before the actual median date and the other half after. A correct' EDD should of course concur with the actual median date of delivery in the population.

The population-based approach to prediction of EDD (Gjessing *et al.* 2007), and to establishment of other categories of fetal biometry charts (Sahota *et al.* 2008, Kagan *et al.* 2009, Verburg *et al.* 2008b, Pexsters *et al.* 2010) seems to be gradually replacing the smaller, conventional studies from earlier decades. Thus, taking advantage of huge databases, into which data of adequate quality from unselected and representative populations have been continuously collected, is an indisputable opportunity.

The unavoidable problems with the traditional selections

In order to construct a traditional, sample-based pregnancy-dating model, it is necessary to recruit selected pregnant women with 'certain' LMP dates. In addition, there must be fetal measurements from each of the GA weeks in the intended range of the study, and in approximately equal numbers from every pregnancy week. To ensure a sufficient number of examinations in each week throughout the intended prediction range, the women must be given additional examinations outside the range of the routine examination. As a result, the selected women and resulting examinations are not very representative of the 'typical' expectant mother at a routine scan; moreover, the prediction model is still LMP-based. This may in part explain the biased predictions of the traditional models (Tunón *et al.* 1998, Salomon *et al.* 2010, Backe and Nakling 2006).

To rectify this problem, a population-based model was developed (Gjessing *et al.* 2007), predicting date of delivery directly. The population data will reflect the local

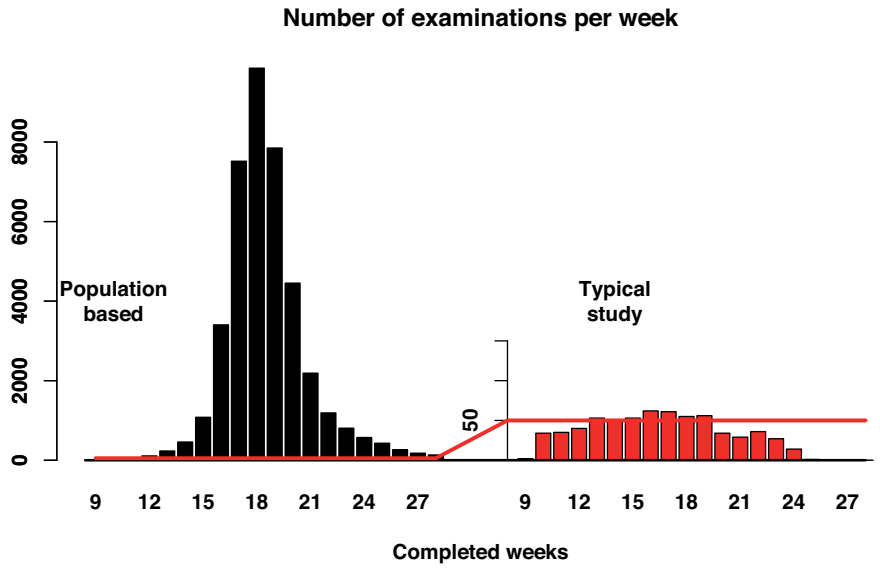


Figure 16. *Figure by H. K. Gjessing, illustrating the inherent difference in the distribution of the included examinations between the population-based model, with examinations concentrated around central weeks, versus the traditional models with a flat sample distribution.*

referral practice and the distribution of the scheduled examinations; with a routine second-trimester examination practice, the majority of the included examinations will be measurements from pregnancy week 17–19, as shown in Figure 16.

The crucial point is that the traditional models were developed on samples with distributions that are different from the populations they are applied to. This causes suboptimal performance. The distribution of the population-based model has a strong concentration of examinations around the central weeks 17–19, similar to the population it is applied to (Figure 16). There are a considerable number of fetuses that are SGA and large for gestational age (LGA) in this central group, and their measurement values spill over to lower and higher values with fewer observations and pull their median remaining time of pregnancy toward the median for 17–19 weeks. Hence, these medians constitute the optimal predictions, paying balanced attention to the average for gestational age (AGA) fetuses and to the spillover of the SGA and LGA fetuses. Being 'trained' on a population distribution, the models will operate in this fashion and produce predictions that are optimal for the population.

In contrast, the sample-based models were developed on data with a flat distribution (Figure 16), effectively paying attention to only the AGA fetus. The significant number of SGA and LGA fetuses from more central dates are erroneously interpreted as too young or too old AGA fetuses, thus producing prediction biases when applied to unselected populations. The effect will be more pronounced the farther away from the central weeks the fetus is examined. In other words, the prediction bias for the sample-based models tends to vary with the fetal size at examination (Figure 11 and 12) (Tunón *et al.* 1998, Backe and Nakling 2006), while the population-based models are aimed at the actual population and avoid these biases (Figure 14).

To ensure lasting optimal predictions, the methods must be re-calibrated when the examination practice and the population distribution change, and future models are not likely to survive unadjusted for 20 years, as did the Trondheim–1984-model.

Fetal biometry

The roles of different measurements

With improved ultrasound machines, fetal structures and organs can be measured at earlier fetal ages, and there is at least one normality curve for every structure (Snijders and Nicolaides 1994, Degani 2001). In first-trimester examinations, the crown-rump length (CRL) is measured and used as the most reliable dating parameter. In addition, the nuchal translucency and other structures may be assessed to evaluate risk of anomalies and chromosomal aberrations. At any GA, a suspected anomaly will trigger the initiation of a more detailed biometric measuring.

In assessment of fetal age during the second-trimester, the preferred measurements and the measurement techniques vary between and within countries, for the most part depending on the preferences of the ultrasound pioneers and on the interests and wishes of the ultrasound environments to achieve standardized fetal age charts. Nevertheless, measurements of the fetal skull (either BPD, or different variants of head circumference (HC)) are basically included, as are measurements of the femur (the FL) — and of the fetal trunk (diameters or circumference of abdomen).

Many investigators have studied the ratio between two different fetal measurements to determine whether a ratio may provide more reliable information about fetal age and growth than the two parameters by themselves (Gottlieb and Galan 2008). There are normality curves for such ratios as well. For example, if growth assessment is the issue, the abdominal circumference (AC) is the denominator (Campbell and Thoms 1977), whereas FL often is included in ratios concerning age (Mul *et al.* 1996) or in the diagnosis of anomalies. More recently, it has been proposed that the ratio between the transverse cerebellar diameter and the AC is better in identifying IUGR than are the other ratios, but the sensitivity remains low (Gottlieb and Galan 2008).

A ratio's potential place is in the assessment of fetal growth — not in pregnancy dating. A ratio may identify asymmetric growth deviations, but of course, the ratio will be normal in a fetus with symmetric IUGR (Degani 2001).

Attempts have been made to increase the accuracy of EDD predictions by combining different fetal measurements in the same regression formula. Hadlock *et al.* (1984) analyzed a number of combinations of 4 measurements and found many of the combinations to be better than using a single parameter (BPD) alone. Ott (1985) also combined the 4 parameters BPD, HC, AC and FL, and indicated that the simple arithmetic average of the GA predicted by each of them resulted in the lowest systematic and random error. Chervenak *et al.* (1998) found HC to be the best parameter, and the predictions were marginally improved by adding AC and FL into the equation. More recent studies have not demonstrated the same convincing effect of combined formulae (Taipale and Hiilesmaa 2001, Salomon *et al.* 2010), and with larger study populations such combinations have become somewhat pointless (Salomon *et al.* 2006b, Verburg *et al.* 2008b)

Various approaches have been tried to evaluate fetal growth in the last half of the pregnancy. One method of evaluation is to repeat the measurements taken at the dating examination and, if indicated, to include additional measurements or ratios. Another method to evaluate growth is by using three-dimensional ultrasound to measure the volume of, for instance, the fetal thigh.

Crown-rump length — CRL

Robinson (1973) was the first to demonstrate that ultrasound measurements of first-trimester fetuses were achievable, and 2 years later a CRL dating curve was published (Robinson and Fleming 1975). This curve has been widely used ever since, despite some inherent weaknesses. The selection criteria were strict and, obviously, the ultrasound machines were relatively old. The publishing of a ‘critical evaluation’ with both the observed CRL values, values derived from regression analysis, and a curve with CRL values after correction for the systematic errors have given rise to a certain confusion among clinicians. Interestingly, the existence of two different CRL curves by Robinson and Fleming is not well known (Koster *et al.* 2008), and hence they are being mixed up, making the measurements incomparable.

Among the systematic errors that were discussed in the second paper, the effect of the beam width in overestimating the CRL measurements was particularly pointed out (Robinson and Fleming 1975, Jago *et al.* 1994).

A recent CRL curve published by Verburg *et al.* (2008b), based on 2079 CRL measurements, gives consistently smaller CRL values than do the conventional curves. Pexsters *et al.* (2010), with 4387 ultrasound examinations, also observed this CRL shortening, particularly in the lower GAs. These findings indicate a significant underestimation of GA of up to 4 days according to Verburg *et al.,* if pregnancies are dated with the old curves before 11 weeks' gestation.

The CRL is measured as the greatest length of the embryo, in a straight line from the cranial to the caudal end of the embryonic body, and in a midline sagittal section of the whole embryo, as shown in Figure 17. CRL may be measured either trans-vaginally or trans-abdominally, and the fetus should typically be horizontal on the screen so that the CRL is measured in the lateral direction with the CRL line in a right angle to the ultrasound beam (Loughna *et al.* 2009). The CRL has been established as the gold standard measurement for fetal age assessment in the first trimester (Pexsters *et al.* 2010, Degani 2001, Chalouhi *et al.* 2011). The 95% confidence interval (CI) of the assessed fetal age is reported to be in the range of only ± 2–3 days before 11 weeks (Degani 2001), while others indicate a wider range with a median error of around ± 6 days (Salomon *et al.* 2010). With increasing fetal age, both flexion and fetal body movements increase, resulting in potential under- or overestimation of GA.

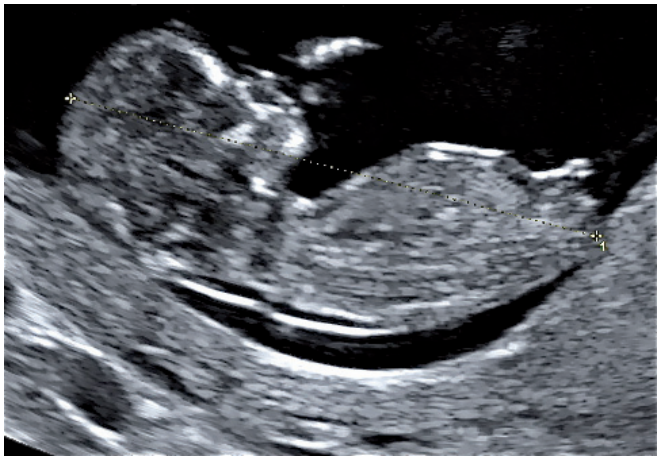


Figure 17. *Measurement of crown-rump length — CRL*

Photo: T.M. Eggebø

The prenatal risk assessment of a first-trimester examination is based on an accurate GA derived from a precise CRL measurement (Snijders *et al.* 1998, Nicolaides 2011a). As a variation in the range of 1–2 days of GA can alter a Down's screening result significantly, it is essential that sonographers and laboratories use the same, standardized dating formula (Loughna *et al.* 2009).

Biparietal diameter — BPD

The potential relation between GA and BPD was described in 1969 (Campbell 1969). Later, the number of studies associated with this single and simple fetal ultrasound measurement have been countless, and it is undoubtedly the most frequently ultrasound-measured fetal parameter.

The BPD is measured in the axial direction, with the ultrasound beam perpendicular to the echoes from the midline brain structures (Figure 18), across the widest part of the skull. Therefore, improvements in ultrasound scanner technology and measurement variations over time influence the BPD to a lesser degree than structures measured in lateral or oblique directions and (Jago *et al.* 1994). The standard measurement plane for BPD is still the one described by Campbell and Thoms (1977), except for what they reported as 'echoes from the third ventricle', which were later identified as being the cavum septi pellucidi (Hadlock *et al.* 1982a).

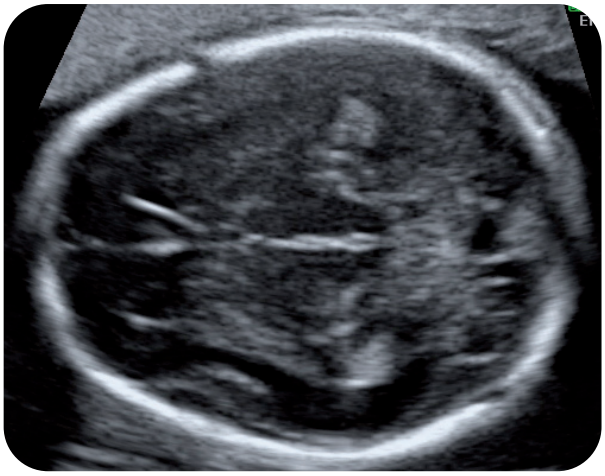


Figure 18. *The correct anatomical plane for measuring the biparietal diameter — BPD*

Both the intra- and inter-observer measurement variation is low; <1 mm (Geirsson 1991), and its geometric features and accuracy makes the BPD measurement easy to obtain, even for rather

unexperienced health workers (Rijken *et al.* 2009). The prediction error of the BPD is described as being around ± 5–7 days, if measured before 20 weeks' gestation (Degani 2001, Taipale and Hiilesmaa 2001, Salomon *et al.* 2010). In fact, a BPD-based prediction in the second-trimester is as accurate as a CRL-based prediction in the first trimester (Taipale and Hiilesmaa 2001, Verburg *et al.* 2008b), and recent studies even show predictive capacity for BPDs 15–30 mm on a par with CRL in late first-trimester (Salomon *et al.* 2010, Chalouhi *et al.* 2011). According to Verburg *et al.* (2008b) the dating of pregnancies could be optimized by using CRL from 20 to 65 mm and BPD from 23 mm onwards. After week 24, ultrasound imaging is not found to give better estimates than the LMP (Verburg *et al.* 2008b).

A well-known drawback of the BPD measurement is that it may be influenced by an unusual fetal head shape — a ‘problem‘ detected more often in the third than in the second trimester, and therefore having more consequences for growth assessment than for dating. The fetal head is usually ovoid in the transaxial view. Brachycephaly means a more rounded head shape, while dolichocephaly is an elongated one. If indicated, the head shape can be objectively assessed by obtaining the occipito-frontal diameter (OFD), representing the long axis, in order to compute the cephalic index (short axis (BPD) / long axis (OFD) x 100) (Hadlock *et al.* 1981). Normal range is said to be within ± 1 standard deviation; 74–83 (Degani 2001, Hadlock *et al.* 1981). The long axis might be used to calculate a ‘corrected’ BPD (Kurtz and Goldberg 1988), but if the normality of one dating parameter is doubted, another should be chosen (Degani 2001, Kurtz and Goldberg 1988), and the FL then seems to be a reliable alternative (Geirsson 1991, Gjessing *et al.* 2007).

An additional remaining issue in an international standardization of the BPD measurements, is where to place the callipers. When the diameter is measured from the outer to the outer contour of the parietal bones, with sound velocity being set at 1540 meter/second (m/s), the actual measurement reflects the true diameter of the fetal head. The outer to inner BPD measurements (leading edge to leading edge) is an inherited tradition from the A-scan time era, when the sound velocity was set at

1600 m/s and the leading edge to leading edge measurement reflected the true diameter. Today, all scanners have an internationally standardized sound velocity of 1540 m/s, and the outer–outer measurement (true diameter) is 3–4 mm larger than the outer–inner. The confusion is increased by the fact that most of the countries that practice outer–inner measurements of isolated BPD measurements, use outer-outer when the BPD is included in a HC measurement (Loughna *et al.* 2009). In pregnancy dating, the same technique as that used to establish the measurement charts should of course be used (Salomon *et al.* 2011).

Head circumference — HC

In Norway, the HC was not used as a fetal measurement until the Bergen–2004 model included a curve with this parameter. There have been discussions on what this measurement really adds to the accuracy of the assessment of fetal age and growth if the fetal head shape is normal (Gjessing and Grøttum 2007). The HC may be computed from a BPD and an OFD measurement, or by using ultrasound equipment with an ellipse measurement capacity. The mathematical problem is that the HC does not represent the circumference of a true ellipse, as the fetal head is rounded posteriorly — and its shape will never be similar to a 'rugby football shape', which has been described as giving a correct HC measurement image (Loughna *et al.* 2009). The ideal measurement plane is the same as the BPD plane shown in Figure 18 (Campbell and Thoms 1977). As with the BPD measurements, there is no international agreement on where to place the callipers to measure the HC (Salomon *et al.* 2011), nor on the formula for an elliptical circumference. The equation is BPD plus OFD multiplied with a factor of π/2 (1.57) according to the BMUS (Loughna *et al.* 2009), or 1.62 according to the ISUOG guidelines (Salomon *et al.* 2011). This is an adaptation to the mean — not the mathematical formula for elliptical circumference. The section through the planum biparietale where we derive the BPD, OFD as well as make our HC measurement, does not have an elliptical shape. If the fetal head is deformed as a consequence of, for example, breech position or a certain pressure

imposed on the fetal head by the uterus, the shortening of the BPD is not necessarily compensated by a corresponding elongation of the OFD. The way we traditionally measure the HC certainly does not make it independent of the head shape.

HC should not be used until late in the first-trimester. Recent studies end up with equal conclusions (Verburg *et al.* 2008a, Chalouhi *et al.* 2011): In first-trimester dating the CRL is the best predictor, followed by the BPD. In second-trimester dating, the BPD seems to remain the measurement of choice (Saltvedt *et al.* 2004), also in newer dating models (Verburg *et al.* 2008b, Gjessing *et al.* 2007), and there are probably no systematic differences in EDD predictions from BPD or HC with examinations up to week 20. The important issue in the first half of the pregnancy is the reliability of the prediction method and the knowhow of the sonographer. Older studies comparing HC and BPD in the predictive capacity of assessing GA were chiefly carried out in the 1980s and compared third-trimester fetal age prediction (Kurtz and Goldberg 1988, Hadlock *et al.* 1982b). Dating examinations in the third-trimester are obsolete and unnecessary when the more reliable first- or second-trimester dating is performed as a routine, but HC measurements are often included in fetal growth formulae for use in third-trimester.

Effectively, the HC depends on 2 diameters where the long axis, the OFD, is measured in the lateral direction and is thus vulnerable to the inferior lateral resolution — and may over time be affected by the beam width factor.

Femur length — FL

The FL has been used for GA assessment since the early 1980s (Hadlock *et al.* 1982c), and is considered almost as reliable as the BPD in second-trimester pregnancy dating (Geirsson 1991, Gjessing *et al.* 2007). However, its accuracy decreases with increasing GA (Degani 2001). Hadlock *et al.* found a variability of ± 9.5 days in GA assessment with examinations between 12 and 23 weeks of pregnancy (1982c). As ossification of the femur diaphysis does not become clearly visible until a GA of approximately 12 weeks, the FL is not a useful parameter in the

first-trimester examinations (Verburg *et al.* 2008a). During the whole pregnancy, the ossification, and hence the FL-measuring, is limited to the femoral diaphysis. Therefore, femur diaphysis length (FDL) is perhaps an anatomically more correct designation (Deter *et al.* 1987), even though FL is by far the most used. The really correct designation would be 'ossified femoral diaphysis length'.

The FL is usually measured with the femoral diaphysis in a longitudinal section, at an angle of less then 45° to the horizontal plane (Goldstein *et al.* 1987, Salomon *et al.* 2011), as shown in Figure 19. The BMUS-guidelines recommend that the femur be imaged as horizontal as possible (Loughna *et al.* 2009). However, this actually increases measurement error, due to the poor lateral resolution and the beam width effect (Jago *et al.* 1994, Verburg *et al.* 2008a). There is reason to believe that the narrowed beam width over time has influenced FL measurements measured in the lateral direction, in the same way as CRLs that are measured shorter. Newer measurement charts report shorter FL measurements than older charts (Longo *et al.* 2004).

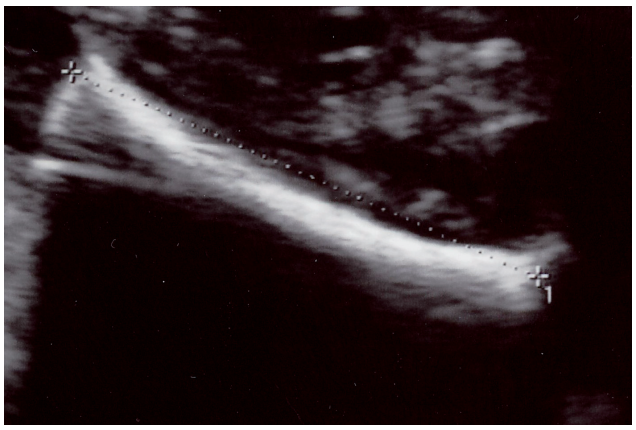


Figure 19. *Measurement of fetal femur length — FL*

There is agreement that the distal femoral epiphysis should not be included in the measurement, but there is varying practice related to whether the measurement should be repeated and, if repeated, whether the mean or the longest of 3 measurements should be recorded. It has been considered correct to generate 3 independent screen images and use the longest measurement (Hadlock *et al.* 1982c, Kurmanavicius *et al.* 1999, Rosati *et al.* 2002), and we have preferred this technique. However, newer guidelines state that provided a technically good image, one single measurement is adequate (Loughna *et al.* 2009).

The length of other long bones, such as the humerus, is considered less accurate than the FL for assessment of dates (Degani 2001). Apart from the dating, FL is also an important parameter in the evaluation of suspected skeletal dysplasia or aneuploidy (Nyberg 2008). However, the finding of an isolated short femur is considered more likely to be related to early fetal growth restriction than to aneuploidy. Whether ethnicity influences FL in second-trimester is still not clear, and the influence of individual maternal and paternal height may be more important than ethnicity (Nyberg 2008).

Abdominal measurements

Measurements such as mean abdominal diameter (MAD) (the mean of the anteroposterior and the transverse abdominal diameter), or abdominal circumference (AC) are widely used in fetal growth equations. They are also routinely included in pregnancy dating examinations, being an important part of a fetal biometry, but they are not in themselves considered to be reliable dating parameters.

The abdominal measurements are measured on a transverse section through the fetal abdomen, with the outline of the abdomen as circular as possible. The spine and the descending aorta are identified posteriorly, the umbilical vein is in the anterior 1/3 of the abdomen and the stomach bubble should be visible in the same plane. On this cross-sectional image, the transverse and the anteroposterior diameter are measured in right angles to each other, from the outer to the outer border of the body outline (Eik-Nes *et al.* 1982b). The AC is measured on the same level, also with the callipers on the outer borders (Loughna *et al.* 2009).

Mathematically, circumference measurements include two diameters that are orthogonal to each other, regardless of whether circumference is being calculated from these diameters or measured directly with ellipse callipers. Independent of the fetal position, such measurements will, in essence, be an approximately equal combination of measurements in the axial and lateral direction.

Standardizing measurement routines on a national level

At the Norwegian consensus conference on obstetric ultrasound in 1986, there was an overall agreement that 'ultrasound examinations in pregnancy must be conducted by personnel with specialized competence — preferably gynecologists and midwives' (Backe and Buhaug 1986). In order to educate midwives in sonography, a formalized education was established in 1997 at the Norwegian University of Science and Technology, physically located at the National Center for Fetal Medicine (NCFM) in Trondheim. Until recently, this was the only teaching unit of its kind in Scandinavia, and it offers a comprehensive sonographic education in obstetric ultrasound. Newly, a similar education has been established in Helsinki, Finland, and a distant learning program is presently being developed at the medical university Karolinska Institute, Stockholm, Sverige.

The value of experienced sonographers and standardized routines to detect congenital anomalies at the second-trimester examination, is reported in several articles from the NCFM environment (Tegnander and Eik-Nes 2006, Brantberg *et al.* 2007, Offerdal *et al.* 2008). The advantages of uniform measurement techniques for reliable pregnancy dating were shown in one of the studies by Tunón *et al.* (1998).

The midwives in Trondheim use internationally standardized measurement techniques for the measurements of the BPD, FL and MAD. Emphasis is on the respective anatomical measurement planes, the setting of the gain on the image and where to place the callipers. Such a practice is important to ensure that the sonographer students later obtain results at their home institution that are uniform on a national level.

Before introducing new measurement routines, a validation ought to be carried out to study if the new measurements actually increase the quality of the EDD predictions, as compared to the results with the established measurement routines.

Technical improvements

Resolution

In two-dimensional ultrasound imaging mode, we scan in a plane that the engineers refer to as the azimuth plane (a terminology adapted from radar), generating the image orientation that is presented on our screen image. In this scan plane, we may measure structures in all directions: axial, lateral or in various oblique directions, as shown in Figure 20. In three-dimensional imaging mode, the elevation direction may also be used for measurements.

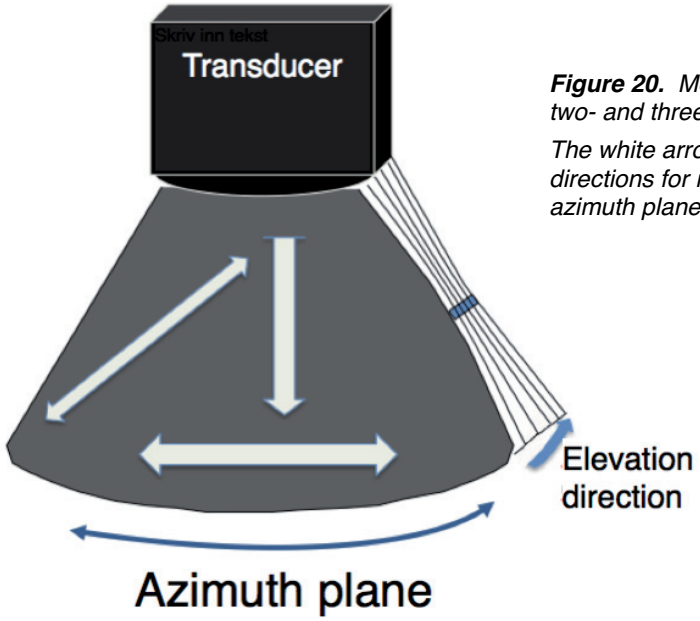


Figure 20. *Measurement directions with two- and three-dimensional ultrasound. The white arrows illustrate the various directions for measurements in the azimuth plane; axial, lateral or oblique.*

Sonographic resolution is defined as the smallest possible distance between two echoes, still enabling us to distinguish them as two separate structures. The resolution in the various directions is dominated by differing physical features. An ultrasound beam that emerges from our transducer travels in the axial/radial direction (most often meaning vertical direction on the screen) into the tissue. This axial resolution is determined by the frequency and bandwidth of the transducer, implying that higher frequencies and shorter pulses give the best axial resolution, but only in the near field. There must be at least one pulse length between 2 echoes in the axial

direction to see them as separate. Deeper-penetrating low frequencies improve imaging in a larger field, but have less axial resolution (DuBose and Hill 1996).

The axial resolution is superior to the lateral resolution. The lateral resolution is determined mainly by the beam width of the ultrasound system, and the wider the beam, the wider a single echo appears to be (Figure 21). There has to be at least one beam width between 2 echoes to see them as separate. (DuBose and Hill 1996).

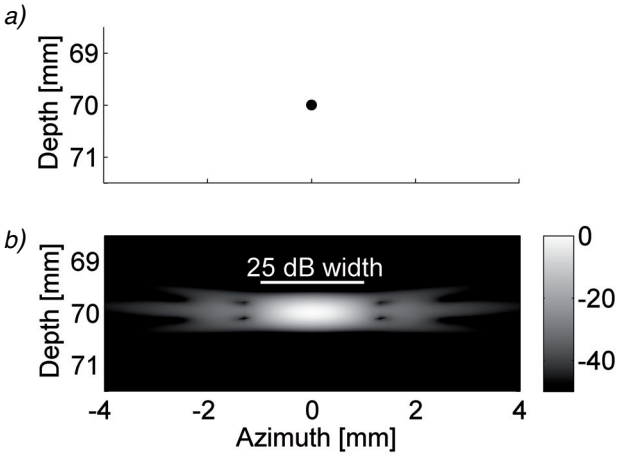


Figure 21. *Ultrasound image of an infinitely small point also referred to as the point-spread function (PSF) of the ultrasound imaging system.*

(a) The point placed in a diagram

(b) Simulated ultrasound screen image of the single point. The 25 dB width of the PSF is indicated.

The fundamental understanding of how axial and lateral resolution affect fetal measurements is extremely important, nevertheless, often neglected by clinicians. Structures measured in the axial direction (BPD) have the least measurement error, while measurements taken transversely to the beam, in directions influenced by the lateral resolution, have greater errors (FL, CRL, OFD), and the lateral demarcation of them often appear somewhat blurred.

These fuzzy echoes make the art of measuring in the lateral direction a rather subjective exercise; the cursors must be placed at the edge of the most definite echo observed and should not include the beam-width artifact — which in fact, is exactly what causes the blurred lateral demarcations shown in Figure 21 b (DuBose and Hill 1996). Therefore, particularly with the old ultrasound machines having wider beam widths, there was always greater inter-observer variation with measurements taken in

the lateral resolution than in the axial resolution. As a consequence, the BPD early became the measurement of choice for reliable second-trimester ultrasound dating of pregnancies. In three-dimensional ultrasound mode, both measurements in the lateral and in the elevation direction will be taken transverse to the axial beam direction and will to equal extents be influenced by the beam-width artifact.

Beam-width narrowing

A reduced beam width has been essential in the development of modern ultrasound machines and has resulted in improved image resolution and better measurement quality. Echoes that originate from the full width of an ultrasound beam are displayed along the centerline of the beam, whether they in reality are located close to it or further (up to half a beam width) away; consequently, up to one beam width may be added to measurements taken in the lateral resolution (Figure 22).

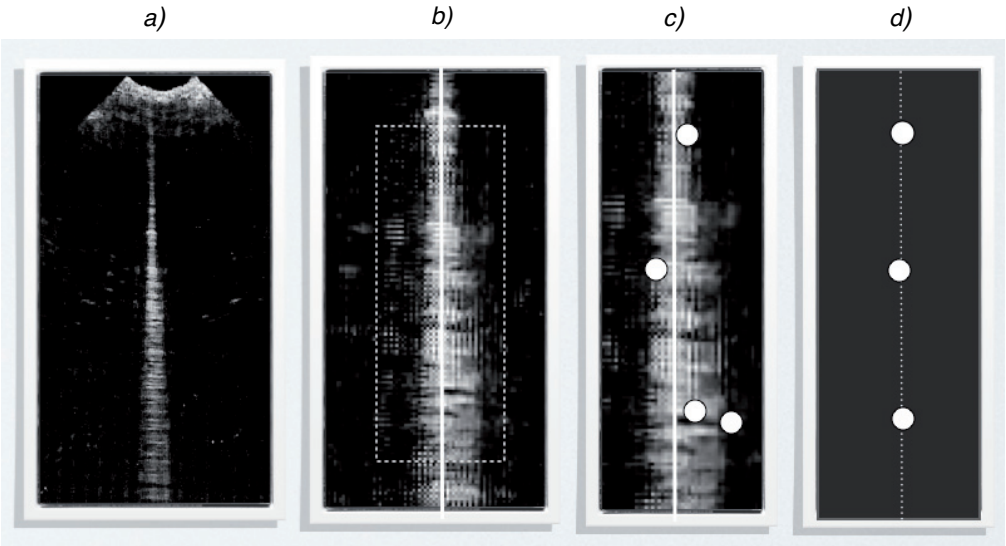


***Figure 22**. Echoes that are 'picked up' by an ultrasound beam are displayed along the imaginary beam centerline on our screen image.*

(a) Ultrasound beam captured with acoustic camera

(b) Ultrasound beam with centerline

(c) Expanded dotted area from b), with centerline and 4 structures (round dots)

(d) Screen image of c) with 3 of the 4 structures picked up by the beam and displayed along the centerline

As shown with the small point in Figure 21, the echoes from fetal structures will also be extended laterally, thus influencing measurements in the lateral direction. For example, with a nearly horizontally measured femur, the echoes from each femoral end are 'picked up' by beams whose centerline is outside the actual ends. The lateral extension effect may amount to 0.5 beam width at each end — one beam width in total — in excess of the true measurement. This beam-width problem was given particular focus by Robinson and Fleming (1975), when they discussed errors in their CRL-curves, and it was raised again by Jago *et al.* (1994), in a study demonstrating that FL was measured significantly longer when measured with an old scanner than with a new one. They discussed how the beam-width narrowing over time might affect measurements in the lateral direction. In more recent years, several authors have mentioned that their new measurement charts included significantly shorter FL and CRL measurements than older curves (Verburg *et al.* 2008a, Verburg *et al.* 2008b, Pexsters *et al.* 2010, Longo *et al.* 2004), but these observations have commonly been attributed to improved ultrasound machines in general, or to different population characteristics, other measurement techniques, or unreliable GAs.

In assessment of fetal age and growth, as well as in anomaly scanning, unbiased fetal measurements are essential. Older dating curves (Robinson and Fleming 1975, Hadlock *et al.* 1982a, Hadlock *et al.* 1984) are still widely used (Saltvedt *et al.* 2004, Koster *et al.* 2008). Using an old measurement chart with a modern ultrasound machine may cause systematic errors in estimations (Jago *et al.* 1994); measurements will be found to be 'too short' and GA will be underestimated (EDD set too late). As discussed previously, such dating errors may have consequences for clinical management and risk assessment.

Aims of the thesis

Correct fetal age assessment is an important part of modern pregnancy care. Ultrasound-based dating from fetal measurements is now the method of choice for reliable estimations of date of delivery. However, the prediction models are shown to vary in predictive quality. While traditional ultrasound-based dating models have been sample-based and the sampled women have been strictly selected, large databases and modern technology have introduced the option of using data from unselected populations as a basis for term prediction.

Moreover, there have been tremendous improvements in ultrasound technology over the last decades. The advances in digital hardware have enabled narrower ultrasound beams, which improves lateral resolution and image quality. Earlier measurement charts were computed from fetal measurements done with ultrasound scanners which are now outdated. One may question if measurements have been influenced by the technical improvements over time and whether the old charts are still reliable.

The aims of the studies were:

Study 1:

To evaluate three sample-based models for prediction of date of delivery on data from the large population-based registry in Trondheim, Norway. We evaluated the predictions from two Norwegian models on 41 343 ultrasound examinations and compared with an established German model (Økland *et al.* 2010).

Study 2:

To validate a new population-based term prediction model and compare the predictions from this model with predictions of date of delivery from the two traditional Norwegian regression models, using data from a clinical database with 9046 routine ultrasound examinations in Stavanger, Norway (Økland *et al.* 2011b).

Study 3:

To confirm the results from the two previous studies of term prediction models in a third population of 23 020 ultrasound examinations performed in Oppland County, Norway. If the previously observed bias could be reproduced, an additional aim was to explore why such a bias occurs with sample-based prediction models (Økland *et al.* 2011c).

Study 4:

To assess the significance of narrowing beam width over time, by exploring potential changes in ultrasound measurements taken in the lateral direction, consequently being potentially influenced by improved lateral resolution. Moreover, to compare beam-width measurements in old and new ultrasound machines to evaluate the extent of the reduced beam-width in modern ultrasound scanners (Økland *et al.* 2011a).

Subjects and methods

Subjects

Studies 1 and 4:

The data for these two studies originated from second-trimester routine fetal ultrasound examinations at the National Center for Fetal Medicine, at St Olavs University Hospital in Trondheim, Norway. As the included women also subsequently delivered at this hospital, complete information about the date of the ultrasound scan and the date of delivery were available. The population was non-selected, coming from a geographically well-defined area. Data were prospectively collected over the period 1987–2005, and more than 30 experienced and formally trained midwives performed the examinations.

Complicated pregnancies related to stillbirths (*n* = 478), diagnosed anomalies (*n* = 1935), multiple pregnancies (*n* = 696) and induction of labor for reasons other than post-term pregnancy (*n* = 4944) were excluded. There were 772 inductions of labor due to post-maturity, and these were not excluded.

In **Study 1,** we included examinations in which the fetal BPD was measured in the range 25–60 mm or the FL was in the range 11–42 mm. Thus, 41 343 examinations in 36 982 pregnancies were analyzed.

In **Study 4,** data from a total of 41 941 examinations in 38 725 pregnancies were included, each with a BPD in the range 35–55 mm, an MAD in the range 32–53 mm, or a GA in the range 113–152 days (corresponding to 16 + 1 to 21 + 5 weeks) at the time of the ultrasound examination. To assess the beam-width narrowing and its impact on measurements in the lateral plane, the data on the FL values from the ultrasound examinations were selected for evaluation. The study material was divided into three approximately equal time-periods; 1987–92 (*n* = 13 354), 1993–98 (*n* = 13503) and 1999–2005 (*n* = 15 084).

Study 2:

The women included in this study had their routine fetal ultrasound examination and subsequent delivery at Stavanger University Hospital, Norway, between January 2001 and November 2006. In accordance with the Norwegian practice of routine ultrasound examinations, the scan was scheduled to take place between 17 and 20 completed weeks according to the LMP and/or an early clinical examination. We included all women with singleton pregnancies, whose fetus had a BPD in the range of 38-60 mm at the routine scan, and who gave birth to a live born child without anomalies after 23 completed weeks.

A total of 10 193 women were initially included. Of these, 1147 were excluded: 704 were excluded because of induction of labor for reasons other than post-term pregnancy, 442 had an elective Cesarean section prior to the start of labor, and one had missing information, leaving 9046 women in the study for analysis. Women induced for post-term pregnancy (n = 301) were not excluded.

Study 3:

The data in this study were collected over a period of 22 years, from 1988 to 2009. They comprise fetal ultrasound examinations performed in Oppland County, Norway, mostly from the two maternity wards in the county, at the Gjøvik and Lillehammer hospitals.

Pregnancies with a fetal BPD in the range of 38–60 mm or an FL in the range of 21–42 mm at the routine ultrasound examination were included. Multiple pregnancies, pregnancies complicated by stillbirth, diagnosed anomalies, induction of labor for reasons other than post-term pregnancy, or elective Cesarean sections were not included. In total, fetal measurements from 23 020 second-trimester examinations were included. Data from women with an available date of LMP (n = 19 131) were used to assess the differences between the LMP-estimated GA at the actual time of the deliveries and the EDD predicted from the BPD measurements with each model.

Methods

The evaluated prediction models and the ultrasound measurement technique are described in the Introduction of this thesis and in each of the papers. In all of the study populations, the Trondheim–1984 model was the dating model used during the study period, and all clinical problems were managed according to this model. As this model predicted date of delivery later than the other models for all BPD measurements, a potential bias in our analyses related to, for example, the post-term inductions, is avoided.

Statistical methods

Studies 1, 2 and 3:

The prediction models to be evaluated were applied to measurements from the ultrasound examinations and to data from the subsequent deliveries.

To correct for the narrowed beam width in newer ultrasound scanners, a correction for the time period that applies to the FL measurements was included for the newer models in Study 1 and 3, where the data collection started in the late 1980s (Gjessing *et al.* 2007). Newer prediction models should not unrestrictedly be applied to older data.

The resulting term predictions were compared with the actual time of delivery, and the disagreement was assessed in terms of the median bias for each model, which reflects the systematic error of the term predictions. Predicting term too early results in a positive bias and an increased rate of apparently post-term pregnancies, while predicting the EDD too late compared with the actual delivery date, gives a negative bias and an apparent decrease in the rate of post-term pregnancies. In addition to for the study population as a whole, the median biases were calculated for subgroups with different fetal ages, since a bias that varies with the fetal size at the time of the ultrasound examination may be missed if only the overall median bias is computed (Gjessing and Grøttum 2007). In Studies 1 and 2, also secondary measures were

computed, *i.e.*, the proportion of births within ± 14 days of the EDD, and the rates of preterm (≥ 24 days before the EDD) and post-term (≥ 14 days after the EDD) deliveries. An altered median bias will affect the rate of deliveries defined as post-term and the proportion of births within ± 14 days more than it will affect the rate of deliveries defined as preterm, due to the left-skewed birth distribution curve. The secondary measures are influenced by both bias and precision.

The bias indicates the systematic error of the predictions, *i.e.*, the calibration of the model and is of relevance when comparing different models based on the same measurement. Consequently, a bias can be removed simply by calibrating the model (Gjessing and Grøttum 2007). Precision is the spread of the residual distribution — the random variation around the median prediction value — independent of the size of the bias. It is important to evaluate precision when assessing the predictive quality of different fetal measurements, for instance BPD versus FL or HC. A lack of precision is more problematic than a bias, because precision is not improved by calibrating the model — one has to improve measurement quality. As an example, even predictions made from LMPs can be made unbiased by using the correct average pregnancy duration, but this does not mean that LMP is as precise as ultrasound in predicting EDD (Gjessing and Grøttum 2007).

The median is the most robust parameter to consider when evaluating pregnancy length and prediction models (Gjessing *et al.* 2007, Salomon *et al.* 2010), and in addition, it may be pedagogically and convincingly communicated as the day by which half of the births have taken place. The alternatives, the mean and the mode, are obviously less robust. The mean is too sensitive to the outliers, mostly represented by the preterm deliveries that skew all birth distribution curves heavily to the left. The earlier a delivery occurs, the more it influences the mean (Gjessing *et al.* 2007). One may try to 'solve' this problem by excluding preterm deliveries from the analyses (Saltvedt *et al.* 2004, Chalouhi *et al.* 2011). Such exclusions may reduce the mean pregnancy length by many days (Saltvedt *et al.* 2004, Gjessing *et al.* 2007).

On the other hand, the mode is too insensitive. Only the very normal births, occurring within a day or two of the mode day, influence this parameter.

95% CI for the bias values, illustrating the uncertainty of the bias estimations, were computed using bootstrapping with 2000 replications. In Study 2, also P-values were included, indicating whether the median biases were statistically significantly different from zero. All analyses were produced in the R statistical programming environment (2010).

Statistical methods, Study 4:

We carried out beam-width measurements on a tissue-mimicking phantom (Figure 23) CIRS model 40, measuring at 6 different depths, using the caliper function of the scanners. The phantom consists of several reflective strings at different depths, enabling measurements of the point-spread function (PSF) of an ultrasound imaging system, as seen in Figure 21. The width of the PSF relates directly to the resolution, and hence the beam width of the scanner, at a specific depth. Thus, the term ‘beam width’ refers to the width of the system’s two-way beam, which is the effective beam resulting from the combination of the transmit and the receive focusing. The focus was optimized for the imaging range of 3–8 cm.

As in clinical measuring, the beam width was defined at the middle grey tone level of the PSF image. This level corresponds to the middle of the dynamic display range that is the range of echo intensities displayed on the screen. A simulated example of such a measurement is shown in Figure 21, where the referred width corresponds to a 25 dB drop in echo intensity compared with that of the white level. To enable comparison of beam widths across different

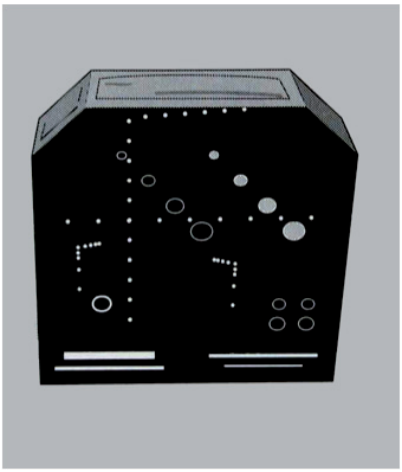


Figure 23. *A tissue-mimicking phantom*

machines, the gain and dynamic display range were adjusted to be as similar as possible in old and new machines.

One operator performed all the phantom measurements and was blinded for the results during the procedure. The sequence of six measurements (at depths 3, 4, 5, 6, 7 and 8 cm) was repeated 25 times with each scanner.

We analyzed the 41 941 FL measurements to evaluate the possible effect of a changing beam width over time. First, median FL values were computed for each day of GA in each of the three time-periods. Similarly, median FL values were computed for each BPD (in mm) and each MAD (in mm) in the three periods. This allowed an assessment of whether median FL values vary over time, independent of fetal size and age. We thus controlled for potential changes in fetal growth pattern or in the time of routine ultrasound examinations. Second, to obtain a summary of the change in median FL values over the three time periods, we analyzed the data with a quantile regression model, using FL as the dependent variable and time-period as the categorical variable. Three separate analyses were done, adjusting for GA, BPD, and MAD, respectively. Third, to obtain a more detailed picture of the change in FL over time, we did the same quantile regression analyses, replacing the three time-period categories with finer categories spanning one year each. Again adjusting linearly for GA, BPD, and MAD, we obtained median FL values for each one-year category, standardized for GA, BPD, and MAD separately.

To analyze the beam-width measurements from the phantom we used a linear mixed-effects regression model. We regressed the measured beam width on the measurement depth and machine generation. Machine generation was used as a covariate with two levels: old and new. Since 25 measurement replications were made for each depth on each of the six ultrasound machines, we controlled for within-machine dependent measurements by adding machine to the regression model as a random effect with six levels. All analyses were produced in the R statistical programming environment (2010).

Results and comments

Study 1

Biases of traditional term prediction models: results from different sample-based models evaluated on 41 343 ultrasound examinations.

Ultrasound Obstet Gynecol 2010; 36: 728–734.

Inherent weaknesses in the traditional models' predictions of date of delivery were found, with biases that varied over each model's measurement range. The predictive quality depended on the fetal size at the time of the ultrasound examination.

Results:

The median biases for the BPD-based predictions varied from -4 to +4 days. There were substantial variations in biases; these variations were found both within each model, varying with the fetal size, as well as between the models, as shown in Figure 25 a–b. The disagreement between the two Norwegian models was mostly in the range of 3–4 days and always >2 days. The German model developed by Hansmann showed a varying bias very similar to Bergen–2004 in the main BPD prediction span.

The overall median biases for the two Norwegian models' BPD-based predictions were: Trondheim–1984, -1.43 days; Bergen–2004, 1.52 days.

The FL-based predictions from the Bergen–2004 model were also evaluated. The prediction bias showed the same variation as the bias of the BPD-based predictions from the same model: increasing bias with increasing fetal size. The within-model variation extended to 5 days, dependent only on the fetal size at the examination (Figure 25 c). The overall median bias was 2.61 days.

Comments:

The traditional prediction models that were evaluated in this study were constructed from relatively small samples of selected women with anticipated normal pregnancies and reliable LMP-data. Such a selected study population differs greatly from the women in the population of routine examinations. The observed varying biases seem

unavoidable when the models were applied outside the normal time span for routine examinations or to women other than those who have been highly selected. For the Trondheim—1984 model, the bias following dating prior to pregnancy week 18 were known from earlier studies (Tunón *et al.* 1998, Kiserud and Rasmussen 1999, Backe and Nakling 2006). However, it was presumed, and by many clinicians taken for granted, that the 'new' Bergen–2004 model had escaped the bias problems, even though the model had not been satisfactorily evaluated. Our study showed that the bias problems were inherent, and indicated that the term predictions from Bergen–2004 were even more biased than those from Trondheim—1984 (Figure 25 a–c).

The selection criteria of the sample-based models are of great consequence in producing the biases, which are of clinical significance. To determine whether population-based prediction models might be more suitable, another population of routine examinations needed to be investigated.

Study 2

A new population-based term prediction model vs. two traditional sample-based models: validation on 9046 ultrasound examinations.

Ultrasound Obstet Gynecol 2011; 37: 207–213.

The overall biases, as well as the biases for the subgroups, were all smaller with the population-based model than with the traditional regression models.

Results:

The two sample-based models, Trondheim–1984 and Bergen–2004, exhibited largely the same biases in this study population, as in the population of Study 1. The biases had the same direction and were of equivalent size, and they showed the same tendency of variation with fetal size both for the BPD-based and the FL-based predictions (Figure 25 a–c). The median biases for these models varied between -3.2 and +4.5 days over the inclusion range for the BPD-based predictions; Trondheim–1984 generally estimated the date of delivery too late and Bergen–2004 too early

(Figure 25 a–b). For the same BPD-value, the difference in the EDD between the two models was in the range of 3–4 days. The secondary measures of our evaluation were, of course, directly influenced by the bias of each model.

The population-based model, Trondheim–2007, also was validated and compared with the two traditional models. The bias of the BPD-based predictions was stable and negligible, essentially within ±1 day (Figure 25 d).

The overall median biases for the models' BPD-based predictions were: Trondheim–1984, -1.57; Bergen–2004, 1.74; and Trondheim–2007, -0.15 days.

The FL-based predictions from Bergen–2004 were compared with those from Trondheim–2007. The bias of the predictions from the traditional model showed the same increasing offset with increasing fetal age as was found in Study 1; a bias amounting to almost 5 days when EDD was determined from longer FLs (Figure 25 c). For the population-based model, the FL-based predictions were stable and with a minimal bias within ±1 day (Figure 25 e).

The overall median biases for the models' FL-based predictions were: Bergen–2004, 1.91; and Trondheim–2007, -0.48 days.

Comments:

The finding of reproducible biases in the term predictions from the two traditional models in two different populations indicate that certain deficiencies are built into these models while they are being constructed. As discussed earlier, the biases are probably due to the selection of healthy women with ‘hyper-normal’ pregnancies, the primary estimation of an LMP implying only an indirect prediction of an EDD, few sonographers performing the measurements, the choice of a fixed and possibly incorrect pregnancy length, and most essential, that the traditional models were developed on other sample distributions than the ones they are being applied to.

The notably smaller biases following the use of the population-based model have several explanations. By estimating the remaining time of pregnancy through a direct

regression of time to delivery on fetal size, this model is independent of the LMP. Interestingly, since the model is calibrated to the correct median remaining time, knowledge of the total length of pregnancy is unnecessary for term prediction, but still essential for estimation of gestational age (Gjessing *et al.* 2010) — the exact opposite of the LMP-based models. This model was based on a large population and measurements were made by many sonographers. This makes the population-model robust against selection bias (Gjessing *et al.* 2007, Salomon *et al.* 2010, Hutchon and Ahmed 2001). Finally, the large sample sizes make it feasible to use, for instance, non-parametric quantile regression, which is more flexible and robust than polynomial regressions (Gjessing *et al.* 2010).

Post-term induction routines should be based on reliable estimations of date of delivery; for a population, a few days of displacement in either direction will under- or overestimate the number of true post-terms, resulting in increased risk for post-mature fetuses not being recognized as such, or iatrogenic, 'unnecessary' inductions of labor in women who have barely passed EDD. Post-term pregnancies are associated with a small, but not negligible increase in risk of mortality, probably after 41 weeks (Hilder *et al.* 1998, Heimstad *et al.* 2006, Wennerholm *et al.* 2009). Therefore, a precise term prediction is an important factor for true age-related risk assessment (Norwitz *et al.* 2007).

The incidence of post-term deliveries (≥294 days of gestation) varies in different studies, but is generally below 4–6% when EDD is calculated with ultrasound (Taipale and Hiilesmaa 2001, Wennerholm *et al.* 2009, Lynch and Zhang 2007, Mandruzzato *et al.* 2010). The gradually adopted policy of earlier post-term inductions (Gülmezoglu *et al.* 2009, Norwitz 2011) implies a two- to threefold increase in induction rates, depending on the chosen cut-off. The clinical problems associated with monitoring and/or inducing labor in false positive post-term pregnant women due to inadequate predictions, put an extra strain on delivery units (Wennerholm *et al.* 2009), independent of the discussions on when to induce in post-term pregnancy.

Study 3

Advantages of the population-based approach to pregnancy dating demonstrated with results from 23 020 ultrasound examinations

Ultrasound Obstet Gynecol 2011. DOI: 10.1002/uog.10081.

The study confirmed the findings of negligible biases in the EDD predictions from the population-based model, while those from the traditional models varied substantially. The biases seem inevitable with the sample-based models, and a simple calibration of the models does not remove the bias problems.

Results:

For the traditional models, once again, the biases varied within and between the models and they were also of equal size and direction to what were found in Study 1 and 2. The median biases for these models varied between -4.2 and +4.8 days over the inclusion range for the BPD-based predictions; Trondheim–1984 generally estimated the date of delivery too late and Bergen–2004 too early (Figure 25 a–b). The population-based model predicted EDD within ±1 day from the actual date of delivery also in this population (Figure 25 d).

The overall median biases for the models' BPD-based predictions were: Trondheim–1984, -0.87; Bergen–2004, 2.22; and Trondheim–2007, 0.40 days.

The FL-based predictions from Bergen–2004 were compared with those from Trondheim–2007. The biases of the predictions from the traditional model showed an equally increasing offset with increasing fetal age as was found in Studies 1 and 2, and a bias of 4.5 days when fetal age was settled from longer FLs (Figure 25 c). For the population-based model, the FL-based predictions were stable and with a minimal bias within ±1 day (Figure 25 e).

The overall median biases for the models' FL-based predictions were: Bergen–2004, 1.72; and Trondheim–2007, -0.40 days.

We found a consistently lower discrepancy between the EDD predictions from the population-based model and the LMP-based GAs at the actual time of the delivery, than between the traditional models' EDD predictions and the LMP-based GAs, as shown in Figure 24.

Table 1 shows the percentage of ongoing pregnancies at 4, 7, 11 and 14 days past the EDD predicted by each model. Depending on the prediction model, there is a considerable difference in the percentage of pregnancies classified as post-term.

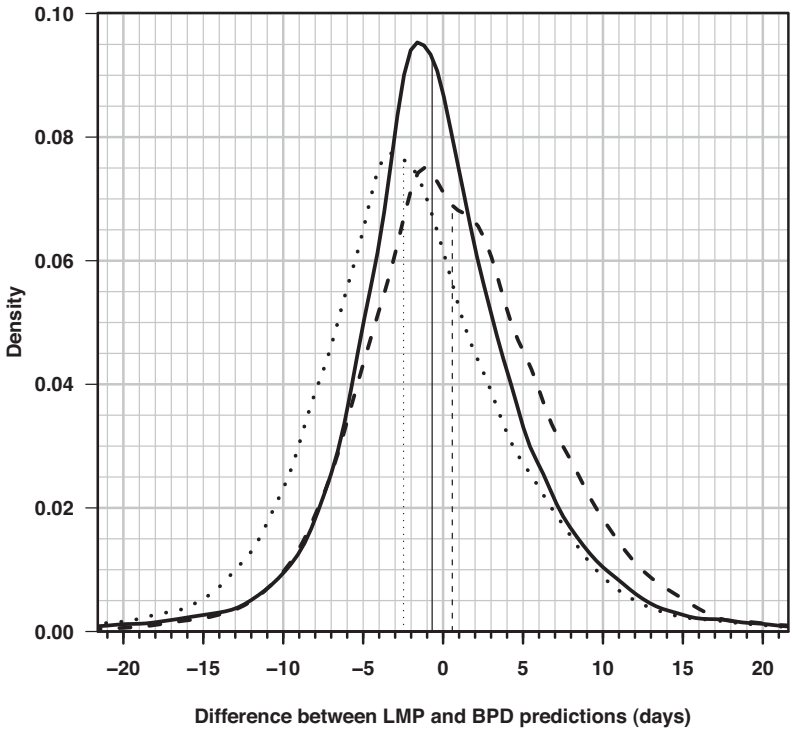


Figure 24. *The differences between the GA at delivery as estimated from the last menstrual period, and the date of delivery as predicted from BPD measurements with the three ultrasound models (Trondheim–1984 (dashed line), Bergen–2004 (dotted line) and Trondheim–2007 (solid line)). The median difference is marked with vertical lines.*

Days past EDD	BPD-based predictions (%)			FL-based predictions (%)	
	Trondheim–1984	Bergen–2004	Trondheim–2007	Bergen–2004	Trondheim–2007
4	31	44	37	42	33
7	20	31	24	28	21
11	9	17	12	15	10
14	4	9	6	8	4

Table 1. *The percentage of still ongoing pregnancies at 4, 7, 11 and 14 days past the date of delivery predicted by each model.*

Comments:

The findings of quite similar, but problematic biases for the traditional models when evaluated on two different populations, and the apparent finding that the population-based model avoided these biases when validated on the Stavanger population, encouraged a study to verify that our results were reproducible. This third study of the term prediction models was on a population from Oppland County, Norway, where there were slightly different routines and other sonographers.

Study 3, together with the two previous studies, comprises a total of 73 400 examinations in three different populations. The studies demonstrate that both sample-based models give systematically biased EDD-predictions, as shown in Figure 25 a–c, while the population-based model does not (Figure 25 d–e).

The basic importance of accurate EDD-predictions in pregnancy care is obvious. The exact knowledge of the GA has an impact on the timing and follow-up of first- and second-trimester examinations and invasive procedures, and on the management of a pregnancy if complications such as IUGR, preterm labor and post-term pregnancy arise (Loughna *et al.* 2009, Dias *et al.* 2011). For the extreme pre-term and post-term

fetuses, a displacement of the EDD in the range of 2 days may have implications for management and survival (Gottlieb and Galan 2008, Pexsters *et al.* 2010).

Comparing the overall median biases of the BPD-based predictions and the BPD bias-curves in Figure 25 a–b for the traditional models, one may wonder why the Trondheim–1984 model has smaller overall median biases than Bergen–2004 in all 3 populations, since both models undoubtedly have fetal size-dependent biases. The difference in median biases are particularly obvious in Study 3: -0.87 for Trondheim–1984, vs 2.22 days for Bergen–2004. This illustrates the importance of considering the biases for different subgroups, not only the overall bias. Moreover, the bias of Trondheim–1984, when used early in the second trimester, has been known for a long time (Tunón *et al.* 1998, Kiserud and Rasmussen 1999, Backe and Nakling 2006), and hence the routine examinations have been scheduled to avoid pregnancy dating until around week 18.

An essential problem is that the traditional models were developed on samples with distributions different from the populations they are applied to. The population-based model was constructed from observations of the actual date of delivery, in order to predict the remaining time of pregnancy and EDD from first- or second-trimester fetal measurements (Gjessing *et al.* 2007, Salomon *et al.* 2010). However, modern pregnancy care requires a precise EDD in the late stages of pregnancy, and knowledge of GA in the early stages. The traditional sample-based models were devised to estimate an LMP, to obtain the GA, from second-trimester fetal measurements and derive the EDD-prediction from this (Hutchon and Ahmed 2001, Gjessing *et al.* 2007) The population-based model estimates the GA as 283 days minus the predicted remaining time of pregnancy. In the reference population (Gjessing *et al.* 2007, Tunón *et al.* 1996), 283 days is the median time from LMP to birth. Since the traditional models are based on estimating the LMP, one might assume that these methods would provide EDDs and GAs that correspond more closely to those computed from the LMP. Interestingly, this is not the case (Figure 24). The EDD predictions from the population-based method correspond more

closely to the GA at delivery as computed from the LMPs of women with reliable LMP-data; this can be seen from the narrower distribution curve of the discrepancy between the LMP- and the ultrasound-based estimates. An overall calibration to remove these median differences (the shifts of the curves away from zero) would not alter the shape of the curves. In other words, even with a recalibration, the population-based predictions would still agree better with the LMP-based predictions.

The better correspondence between ultrasound- and LMP-based estimates has immediate clinical consequences. First, it is beneficial for scheduling examinations and second, it reduces the pregnant women's concern, which sometimes appears if the ultrasound- and LMP-based EDDs differ substantially. Thirdly, it reduces the risk of erroneous dating for fetuses with growth velocity below or above the average (Verburg *et al.* 2008b). The new population-based method is thus better adapted to the actual target population than are the sample-based methods, and will better predict the date of delivery for fetuses with an early, true intrauterine growth restriction (IUGR). However, identifying early IUGR fetuses cannot be done from one single ultrasound examination, irrespective of prediction model. Any significant difference between reliable LMP-based and ultrasound-based EDD-dates indicates a need for further evaluations (Verburg *et al.* 2008b, Tunón *et al.* 1999b, Nguyen *et al.* 2000).

As shown in Table 1, the choice of dating model has a strong impact on post-term induction rates, regardless of the number of days past EDD that are defined as the recommended time for post-term induction. A bias of 2 days in either direction has significant consequences and results in a varying post-term rate between the models. Too early prediction of EDD results in a substantial, yet often ignored, increase in pregnancies wrongly identified as post-term. A prerequisite for comparison of induction routines is unbiased and uniform EDD-predictions with comparable post-term rates (Zeitlin *et al.* 2007). As also indicated in the table, the rates of still ongoing pregnancies are nearly halved for all the models from day 7 to day 11 after EDD.

When considering Studies 1, 2 and 3 together, as has been done in Figure 25 a–e, there is a convincing correspondence between the results. The median biases for each model were calculated from BPD- and FL-based EDD predictions, respectively, in each of the 3 different populations and compared in the same figures. For the population-based model, an evaluation on the population from Trondheim is also included in Figure 25 d and 25 e for comparative purposes (Gjessing *et al.* 2007), although this model was not part of the evaluations in Study 1. The similarity of the 3 individual BPD/FL bias curves for each of the models illustrates that the total population of 73,400 examinations from 3 different locations in Norway were more comparable than we first had reason to believe.

Figure 25 a. Bias curves from BPD-based predictions with Trondheim–1984

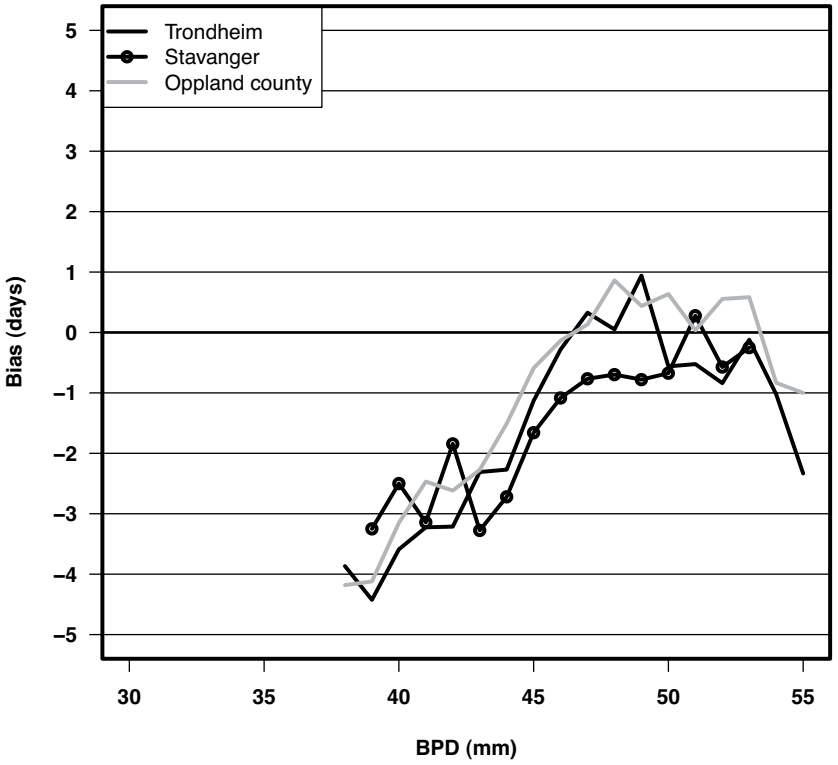


Figure 25*. The median biases for each model, calculated from BPD- and FL-based predictions in the study populations from Trondheim (Study 1), Stavanger (Study 2) and Oppland County (Study 3)*

Figure 25 b. Bias curves from BPD-based predictions with Bergen–2004

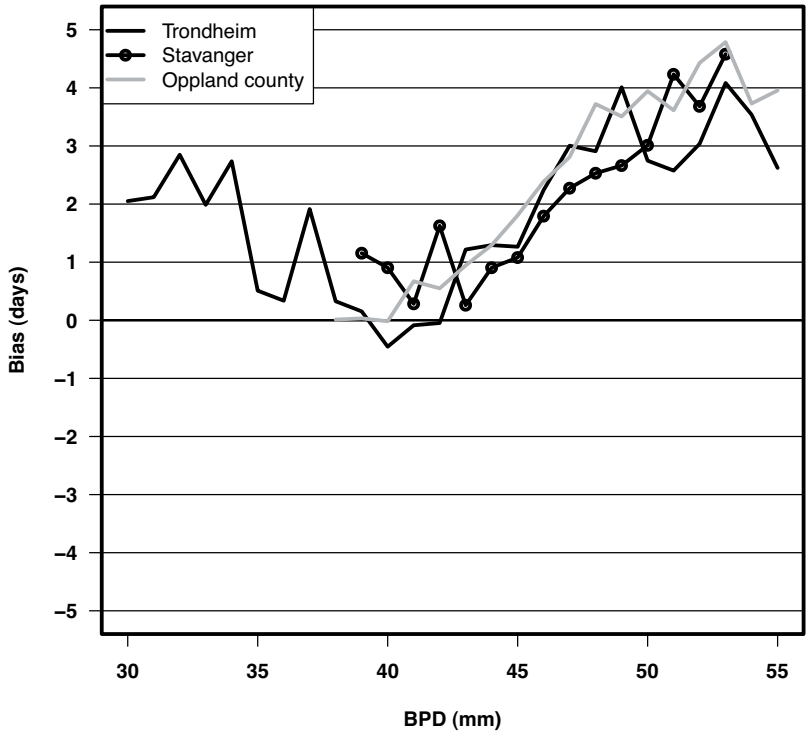


Figure 25 c. Bias curves from FL-based predictions with Bergen–2004

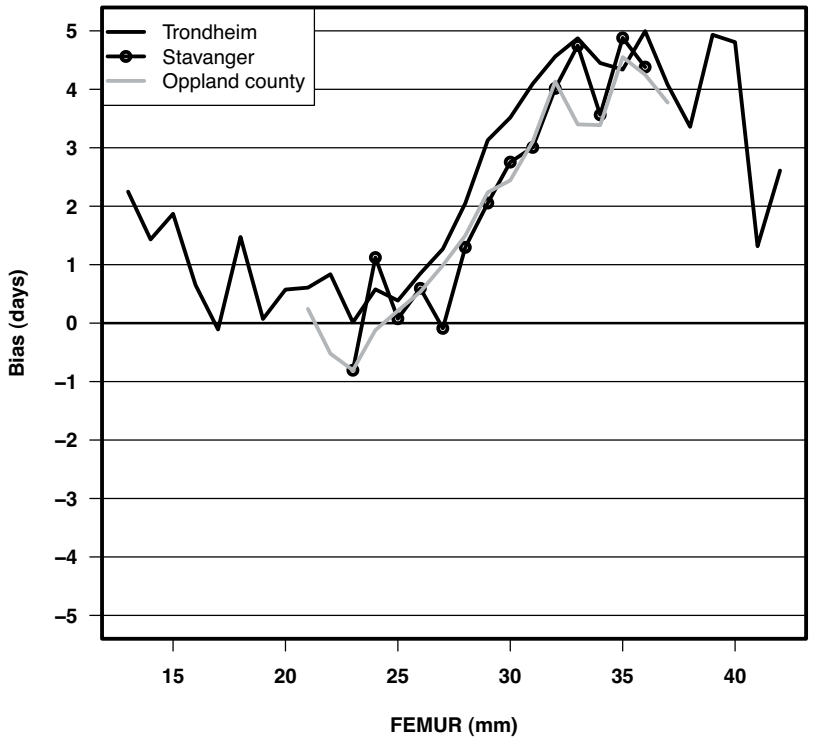


Figure 25 d. Bias curves from BPD-based predictions with Trondheim–2007

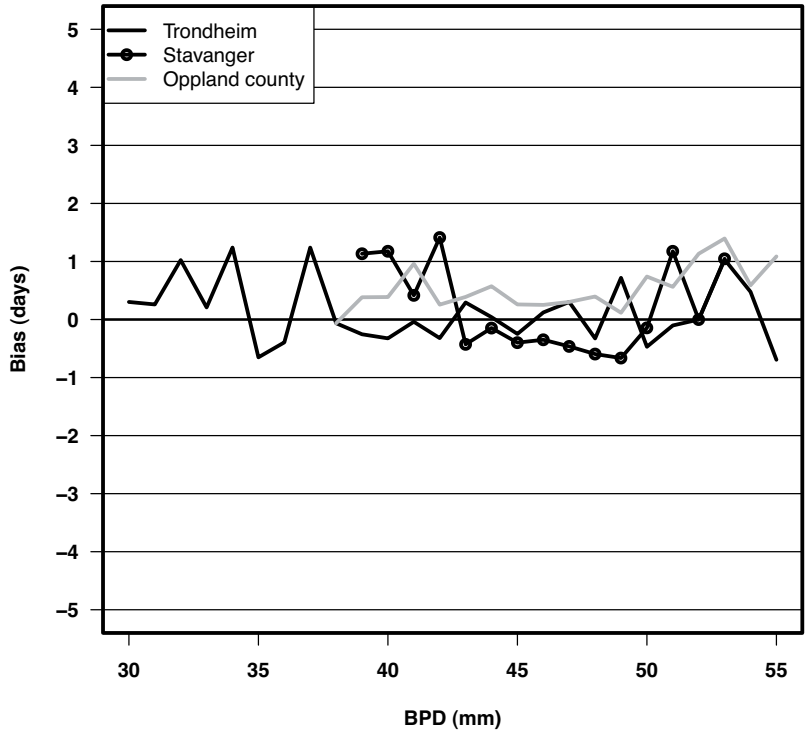
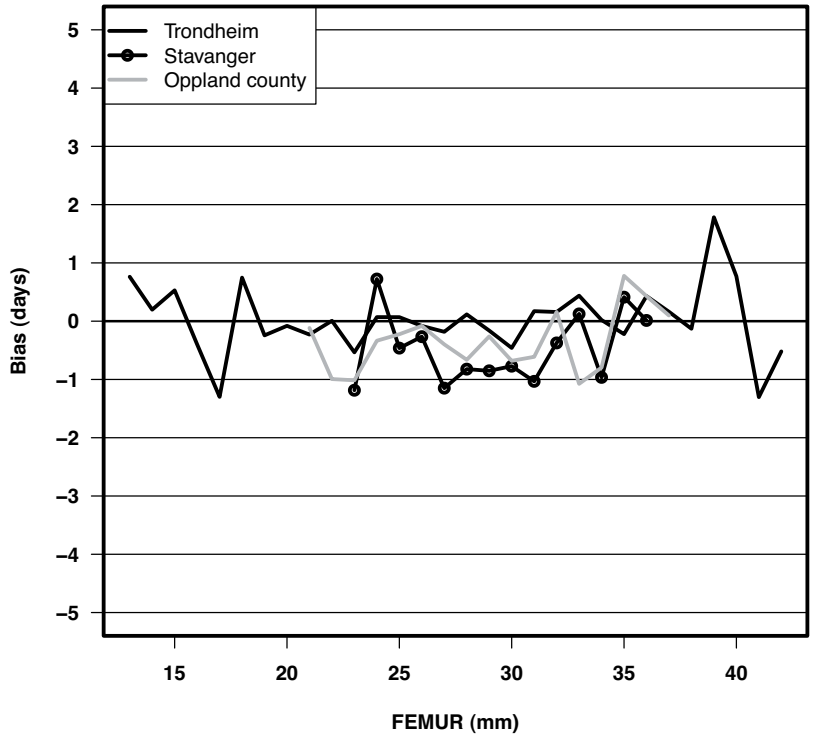


Figure 25 e. Bias curves from FL-based predictions with Trondheim–2007



The midwives who performed the ultrasound examinations at the different hospitals in the presented studies, have all been trained at the NCFM in Trondheim. It is evident that they take their knowledge with them and carry on measuring in the same way when their training is finished. The bias curves in Figures 25 a–e show that there was practically no difference in the resulting median bias for each of the models, no matter where the dating examinations were performed. This is an obvious example of the benefits of education and standardization of routines on a national level.

The fact that Bergen–2004 uses a pregnancy length of 282 days to predict EDD from the estimated LMP, while Trondheim–2007 uses a pregnancy length of 283 days to estimate the GA, has lead to a misunderstanding that this single day is the only difference between the two models, and that adding 1 day to the EDD predictions from Bergen–2004 would do away with the bias. The overall median bias of the predictions from Bergen–2004 are 1.52, 1.74 and 2.22 days for the BPD-based predictions in our 3 populations, and 2.61, 1.91 and 1.72 days for the FL-based. Thus, the needed correction factor is closer to 2 than to 1 day. Furthermore, as shown in Figure 25 and 26, a systematic adjustment is not sufficient. Provided that only the systematic bias resulting from an insufficient calibration of the traditional models was the problem, correcting with a constant value (*e.g.* reconsidering the standard pregnancy length) could potentially eliminate the overall bias (Gjessing and Grøttum 2007). This would correspond to shifting the bias curves in Figures 25 up or down along the y-axis until the median bias is zero. However, the slope of the curves would remain. Particularly for Bergen–2004, since both the BPD and the FL curves slope upward to the right, a correct overall calibration would result in EDD-predictions that are too late for the small fetuses and too early for the large ones. Thus, a simple calibration improving the overall bias would have unfortunate consequences.

A tentative calibration trial of a biased prediction model is schematically illustrated in Figure 26 a-c. The fetal size-dependent biases and the sloping bias curves cannot be corrected with a single constant value, as illustrated in Figure 26 a. A calibration that only incorporates a parallel displacement of the sloping curves upwards or downwards, as shown in Figure 26 b, results in the same absolute error, and the prediction accuracy will still be deficient over the inclusion range. In fact, the optimal and essential calibration is to use population data in an overall calibration, in order to remove the biases over the whole range of measurement values, *i.e.*, to apply a population-based approach to prediction of EDD (Figure 26 c).

Figure 26 a

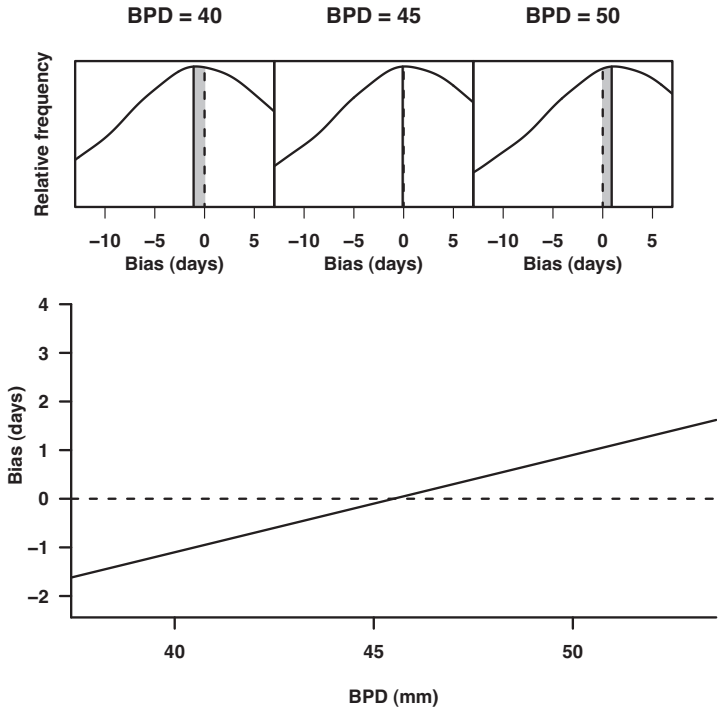


Figure 26; *by H. K. Gjessing. A schematic illustration of a potential trial of a calibration of a biased term prediction model. (a) Bias curve showing offset and a bias that varies with the size of the fetus (b) A tryout of calibration improves the overall median bias, but the sloping curve results in biases of different size and directions for the EDD predictions from the small and big fetuses (c) The essential calibration with the use of population data*

Figure 26 b

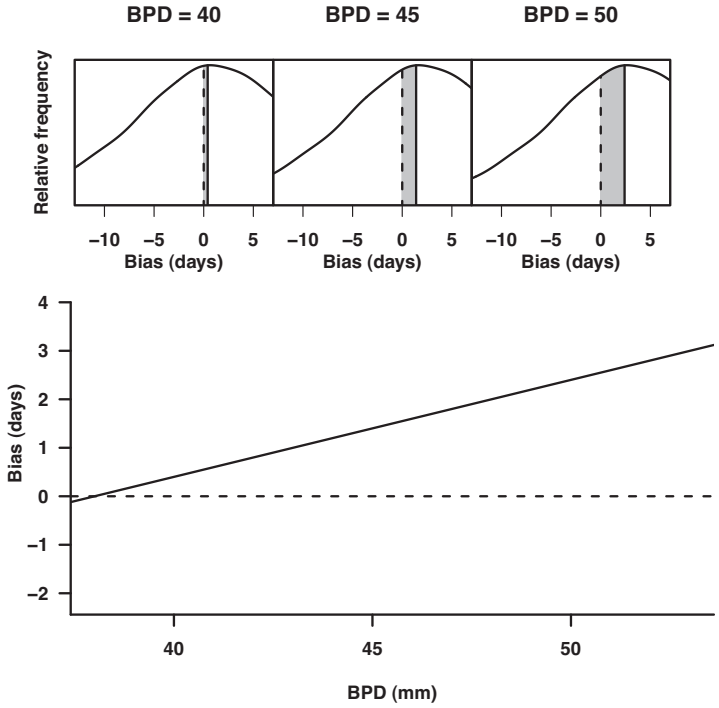
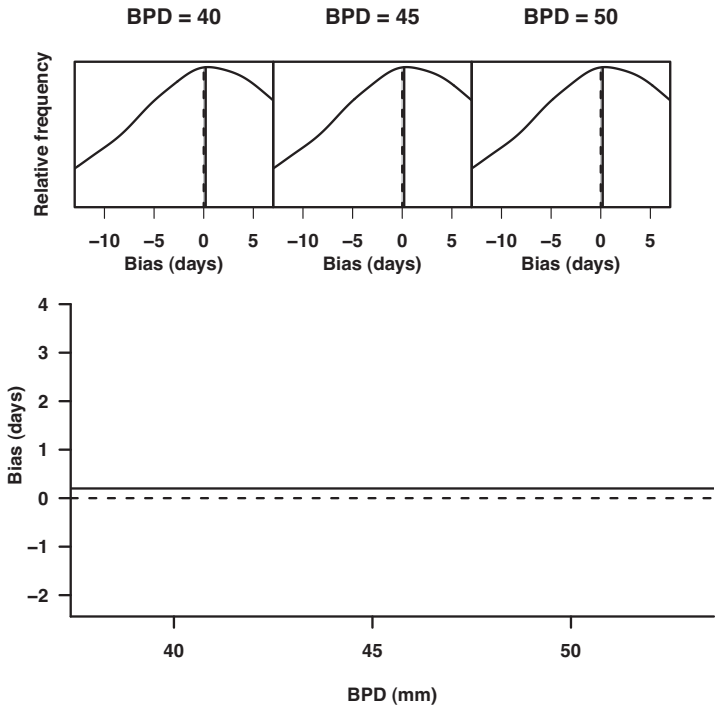


Figure 26 c



Study 4

Narrowed beam width in newer ultrasound machines shortens measurements in the lateral direction: fetal measurement charts may be obsolete

Ultrasound Obstet Gynecol 2011; 38: 82–87.

Technical improvements in modern ultrasound machines that have reduced the beam width affect fetal measurements in the lateral direction.

Results:

Time was a significant covariate. At the same GA, the median FL measurement was 1.15 mm (95% CI, 1.08–1.23) shorter ($P < 0.005$) in the third time-period (1999–2005) than in the first (1987–92). The regressions of the FL measurements on the BPD measurements demonstrated that for the same BPD, the median FL measurement was also significantly shorter: 0.98 (95% CI, 0.93–1.04) mm ($P < 0.005$) shorter in the third time-period than in the first. Regressions of the FL measurements on the MAD measurements showed corresponding shifts, but on a smaller scale: a minor shortening of 0.59 (95% CI, 0.54–0.63) mm ($P < 0.005$) from the first to the third time period. For all the measurements, there were only small differences between the first and the second (1993–98) time period.

Figure 27 shows the estimation of the changes in median FL for each year of the study period, adjusted linearly and standardized for GA, BPD and MAD, respectively. The figure illustrates the main trend and reflects the periods of major replacement of ultrasound machines in the department.

For the beam-width measurements on the phantom, the overall median beam width was 1.08 (95% CI, 0.50–1.65) mm ($P = 0.006$) mm narrower with the new machines than with the old machines. For both scanner generations the measurements increased significantly with increasing depths. Figure 28 shows the results for each of the six machines in box plots.

Figure 27. *Changes in median femur length for every one-year period between 1987 and 2005 adjusted linearly and standardized for gestational age (unbroken line), biparietal diameter (broken line), and mean abdominal diameter (dotted line), respectively*.

Comments:

Over the years, various technical improvements have resulted in narrower ultrasound beams. In this study, the principal reason for obtaining a shorter FL measurement in the last time-period compared with the first and second time-periods, was probably a result of the transition from analog to digital beam formation hardware. In the middle/ late 1990's, digital processing hardware (such as analog to digital converters, application specific integrated circuits (ASICs) and digital signal processors) was included in the scanners. This allowed for much more flexible and precise control of the beam formation process resulting in narrower beam width throughout the whole ultrasound imaging sector. Additionally, an increased system channel number enabled larger apertures, generating a narrower beam. The increased processing

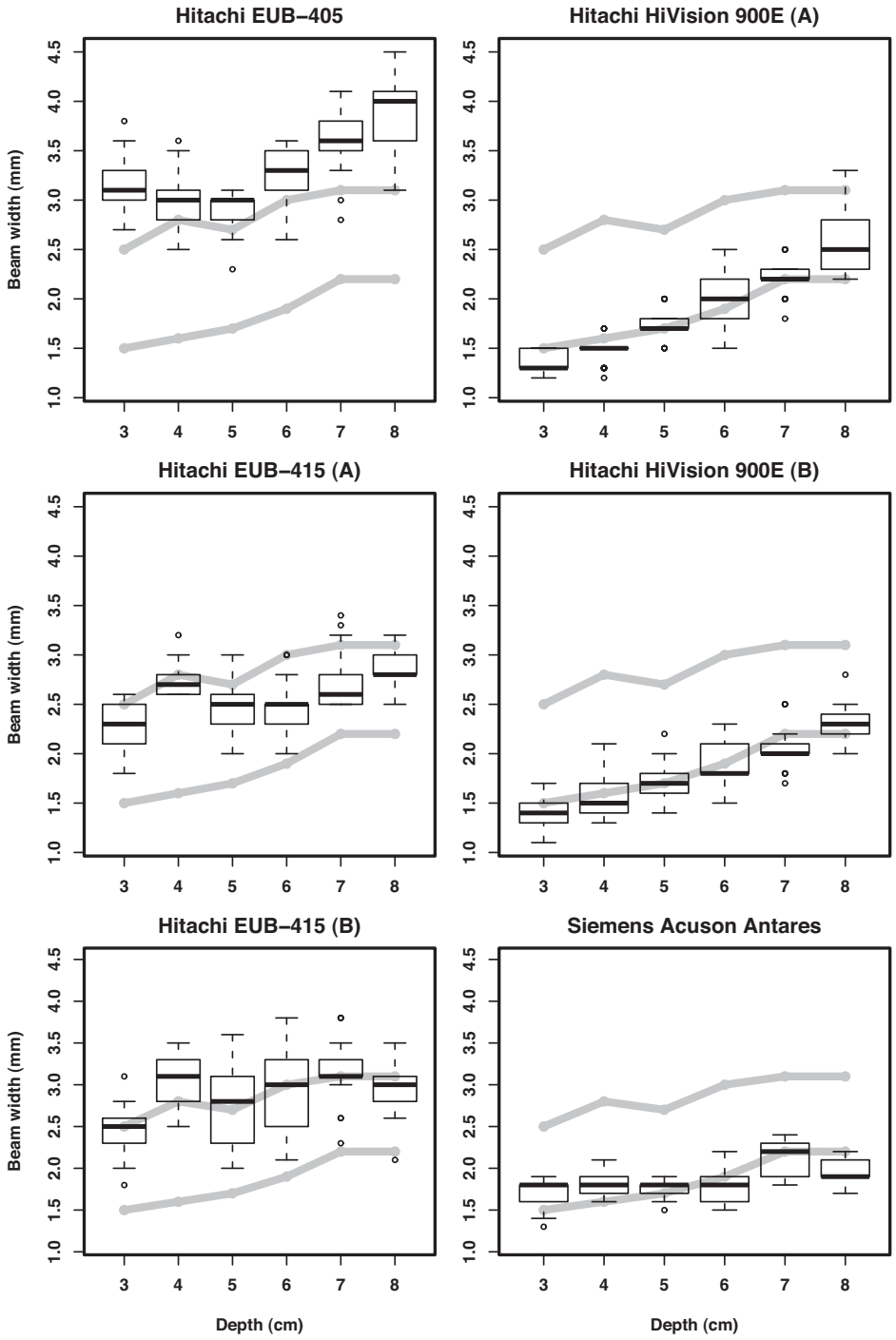


Figure 28. *Box plots illustrating the results of the beam width measurements for each of the six ultrasound machines. The right panel shows the results for the new machines. The grey background curves represent the median beam widths at the various measurement depths for the new (lower curve) and the old (upper curve) scanners.*

power also allowed a high scan line density. Spacing the scan lines too far apart decreases the lateral resolution by more than that determined by the beam width alone.

As the echoes from the lateral ends of structures measured in a direction transverse to the direction of the beam will be extended further laterally, measurements in the lateral direction will be influenced. As shown schematically in Figure 29, with a nearly horizontally measured femur, the echoes from each femoral end are 'picked up' by beams whose centerline is outside the actual ends. Up to one beam width may be added to the true FL measurement when the femur is measured horizontally orientated — as is in fact the FL measurement technique recommended in both the ISUOG and the BMUS guidelines (Salomon *et al.* 2011, Loughna *et al.* 2009). For a femur at 45° the overestimation is reduced to one beam width multiplied by cosine 45°; 0.7 beam width in total, that is 0.35 at each end (Jago *et al.* 1994). Therefore, one should aim at a diagonal rather than a horizontal measurement of the FL, in order to reduce the negative beam-width effect and the guidelines for the two ultrasound societies ought to be changed accordingly.

A potential variation in a certain fetal measurement, for instance FL or BPD for the same gestational age, might be regarded as resulting from a possibly altered fetal growth pattern over time. On the contrary, the relation between the various parameters, *e.g.* how FL relates to BPD/MAD over time indicates how the fetal geometry is being influenced; *i.e.*, how the measurements vary over time, independent of fetal size and age.

Fetal measurements are essential for assessing fetal age and growth, and in fetal anomaly scanning. The measured parameter is related to a measurement chart that is optimized for the purpose of the examination. Curves from the 1980's and 1990's are still widely used as reference charts (Saltvedt *et al.* 2004), despite the fact that charts from the last decade typically are constructed from significantly shorter FL- and CRL-measurements (Pexsters *et al.* 2010, Longo *et al.* 2004, Verburg *et al.*

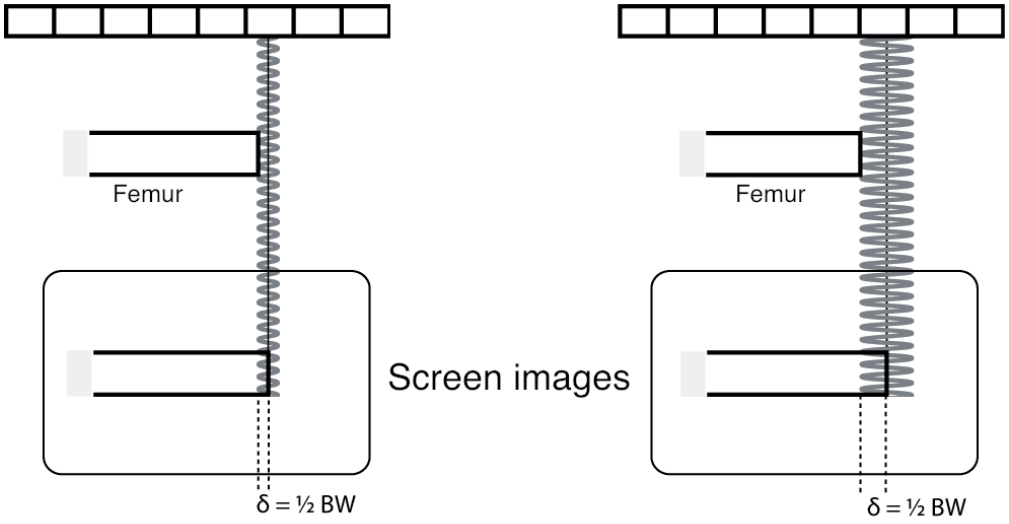


Figure 29. *Schematic illustration of the narrow beam width (BW) in new machines and of the wide BW in old machines, showing how the width of the ultrasound beam influences the lateral extension of an object.*

2008b). It is important to be aware of the technical development and the possible consequences it may have for ultrasound equipment in clinical use. Systematic errors, particularly in the first-trimester assessment of GA, may give dating errors of up to half a week (Pretorius *et al.* 1984, Pexsters *et al.* 2010, Loughna *et al.* 2009) and result in erroneous risk estimations in prenatal screening (Loughna *et al.* 2009, Koster *et al.* 2008). With the use of a modern scanner and an old dating chart, measurements in the lateral direction, such as FL and CRL, will be considered 'too short' (Jago *et al.* 1994), fetal age is then underestimated and the predicted date of delivery is set too late, which of course may have clinical consequences.

Summary and conclusions

Correct fetal age assessment is essential in modern pregnancy care. Ultrasound-based dating from fetal measurements is the method of choice for reliable estimations of date of delivery, and in Norway the dating takes place at the routine ultrasound examination around pregnancy weeks 17–19. The fetal BPD and FL have traditionally been used for term prediction. However, the models for ultrasound-based EDD predictions have been shown to vary in quality. While the traditional ultrasound-based dating models have been sample-based and the women included in the samples have been strictly selected, large databases and modern technology have now introduced the option of using data from unselected populations.

There have been great improvements in ultrasound technology over the last decades. Narrower ultrasound beams have improved lateral resolution and image quality. As old measurement charts were computed from fetal measurements carried out on ultrasound scanners of completely different generations, one may question the measurements made over time and whether the old charts are still reliable when they are used with measurements from modern machines.

The aims of these studies were to evaluate 2 sample-based models for prediction of date of delivery (Study 1), and to compare the predictions from these models with predictions from a new population-based term prediction model (Studies 2 and 3), on data from 3 large Norwegian databases. Study 1 was done with data from 41 343 routine examinations in Trondheim, Study 2 with 9046 examinations from Stavanger, and Study 3 with 23 020 ultrasound examinations performed in Oppland County.

In Study 4 we used the data from routine examinations in Trondheim to assess the significance of narrowing beam width over time. We explored potential changes in measurements of fetal structures (FL) that were measured in the lateral direction. These may be influenced by the improved lateral resolution in newer ultrasound machines. Moreover, we compared beam-width measurements in old and new scanners to evaluate the extent of the reduced beam width in modern machines.

In Study 1, we found inherent weaknesses in the traditional models' EDD predictions with biases that varied over each model's measurement range. The predictive quality of the 2 models was dependent on the fetal size at the time of the examination.

The results in Study 1 were confirmed in Study 2 for the sample-based models' predictions of EDD. The overall biases, as well as the biases for the subgroups, were all smaller with the new population-based model than with the traditional sample-based models. The predictions of the population-based model were stable throughout the inclusion range, and the median bias was largely within ± 1 day.

Study 3 confirmed the findings of negligible biases in the EDD predictions from the population-based model, while those from the traditional models varied substantially. For each of these 2 models, the biases were of similar size and direction to those found in Studies 1 and 2: One model predicted term too early and the other one too late. The biases seem inevitable with the sample-based models, even if tentatively calibrated. The EDD predictions from the population-based model were reliable also in this population, and hence, the population-based model should be preferred as the method of term prediction.

Study 4 showed that the technical improvements in modern ultrasound machines have resulted in a narrowed beam width that affects fetal measurements in the lateral direction. Consequently, the structures appear shorter. With the use of a modern scanner and an old dating chart, measurements in the lateral direction, such as FL and CRL, will be considered ‘too short’. This will underestimate fetal age.

Reliable dating is essential for adequate pregnancy care and is a prerequisite for assessment of potential fetal growth abnormalities, handling of extremely preterm deliveries and in discussions concerning when to induce labour in post-term pregnancies. There is need for a standardization to determine which fetal structures should be measured and which dating chart should be used. The education of sonographers is essential in order to maintain the quality of routine examinations.

Future aspects

Although ultrasound-based pregnancy dating has been a basic part of the pregnancy care for the last decades, there is still a need to inquire whether the prediction models in use perform satisfactorily in the populations to which they are applied. Precise assessment of GA or estimations of date of delivery are obvious prerequisites when calculating the risk of chromosomal aberrations in the first trimester or discussing post-term induction guidelines in the third trimester. Between these extremities there are a variety of situations in which the date of delivery is essential to the management of the pregnancy or of the newborn infant.

An aim of this thesis has been to demonstrate that we should use population data, not only in predictions of date of delivery, but in the construction of all kinds of normality curves. We have the technology, and a systematic collection of data into electronic databases definitely represents the times to come.

These 4 studies are in themselves examples of the importance of a systematic collection of data. The enthusiasts who established the databases that were used in the studies, were ahead of their time; there are still birth departments in Norway that use pen and paper protocols to register ultrasound findings, EDDs and essential data surrounding the birth of a baby.

The advantages of population-based models are their flexibility. They can be continuously updated with new data and if desirable, older data may equally easily be removed if examination routines, measurement techniques, or other time-dependent adaptations bring about reasonable obsoletion of the oldest data. The possibility of a constant quality assessment of the predictions, by immediate population evaluation, is another advantage.

In the near future, a first-trimester ultrasound examination will probably become a part of public health care, also in Norway. So far, this has not been a politically correct option, and consequently around half of the women seek such examinations

outside the public pregnancy care system. If first-trimester examinations are introduced and recommended in national guidelines, the future routine ultrasound dating will probably take place around week 12–14 instead of in week 17–19. The Norwegian dating models in use by now are aimed at second-trimester dating, but the population-based model may easily be expanded with first-trimester data.

Dating of multiple pregnancies is mostly done with dating models primarily constructed for singletons. There are often differences in fetal size between the fetuses, even as early as in the first trimester in spite of the fact that they are conceived simultaneously. It is therefore considered most reliable to date from the largest fetus. However, multiple pregnancies have different birth distribution than singletons. It is obviously difficult to obtain 'populations' of triplets and quadruplets, but a reliable dating and growth model for twins is needed.

Quality assessment is basic in all aspects of health care. If we do not evaluate what we do and how we do it, we cannot improve our achievements. When it comes to pregnancy dating models, other Nordic and European countries still use a variety of old prediction models. In fact, many of these models were constructed from measurements from fetuses in the same generation as the women who today are mothers-to-be — or even soon becoming grandmothers. Few medical facts and methods remain unchanged truths over such a long time, and probably several of these models have become obsolete. Therefore, there is a need for evaluations of predictive capacity in today's populations.

An overall standardization of methods, that is, which parameters to measure, the measurement charts, measurement techniques and not least the education of sonographers, is needed to maintain prediction quality. Obstetric quality indicators and important perinatal outcome measures can thus be compared across regions, and on national and international levels.

References

Altman DG and Chitty LS. Design and analysis of studies to derive charts of fetal size. *Ultrasound Obstet Gynecol* 1993; **3**: 378-384.

Altman DG and Chitty LS. Charts of fetal size: 1. Methodology. *Br J Obstet Gynaecol* 1994; **101**: 29-34.

Altman DG and Chitty LS. New charts for ultrasound dating of pregnancy. *Ultrasound Obstet Gynecol* 1997; **10**: 174-191.

Backe B and Buhaug H. Bruk av ultralyd i svangerskapet. NIS; SINTEF-gruppen. Konsensuskonferansen 27-29/9 1986; Trondheim, Norway.

Backe B and Nakling J. Term prediction with ultrasound: evaluation of a new dating curve for biparietal diameter measurements. *Acta Obstet Gynecol Scand* 2006; **85**: 156-159.

Baird DD, Wilcox AJ, Weinberg CR, Kamel F, McConnaughey DR, Musey PI and Collins DC. Preimplantation hormonal differences between the conception and non-conception menstrual cycles of 32 normal women. *Human Reproduction* 1997; **12**: 2607-2613.

Bakketeig LS, Eik-Nes SH, Jacobsen G, Ulstein MK, Brodtkorb CJ, Balstad P, Eriksen BC and Jørgensen NP. Randomised controlled trial of ultrasonographic screening in pregnancy. *Lancet* 1984; **2**: 207-211.

Baschat AA. Doppler application in the delivery timing of the preterm growth-restricted fetus: another step in the right direction. *Ultrasound Obstet Gynecol* 2004; **23**: 111-118.

Baskett TF and Nagele F. Naegele's rule: a reappraisal. *Bjog* 2000; **107**: 1433-1435.

Beckmann P. In Book *A History of Pi*. The Golem Press: New York, 1971.

Bennett KA, Crane JM, O'Shea P, Lacelle J, Hutchens D and Copel JA. First trimester ultrasound screening is effective in reducing postterm labor induction rates: a randomized controlled trial. *Am J Obstet Gynecol* 2004; **190**: 1077-1081.

Bergsjø P and Brodtkorb C. Ultrasonic fetal cephalometry in pre-eclampsia. *Acta Obstet Gynecol Scand* 1973; **52**: 249-251.

Bergsjø P, Denman DW, Hoffman HJ and Meirik O. Duration of human singleton pregnancy. A population-based study. *Acta Obstet Gynecol Scand* 1990; **69**: 197-207.

Blondel B, Morin I, Platt RW, Kramer MS, Usher R and Breart G. Algorithms for combining menstrual and ultrasound estimates of gestational age: consequences for rates of preterm and postterm birth. *Bjog* 2002; **109**: 718-720.

Boerhaave H. Praelectiones Academicae in Proprias Institutiones Rei Medicae. Haller A (ed). J. Wetstenium: Amsterdam, 1744; Page 437.

Bottomley C and Bourne T. Dating and growth in the first trimester. *Best Pract Res Clin Obstet Gynaecol* 2009; **23**: 439-452.

Boyce A, Mayaux MJ and Schwartz D. Classical and "true" gestational postmaturity. *Am J Obstet Gynecol* 1976; **125**: 911-914.

Brantberg A, Blaas H-GK, Haugen SE and Eik-Nes SH. Esophageal obstruction — prenatal detection rate and outcome. *Ultrasound Obstet Gynecol* 2007; **30**: 180-187.

Bukowski R, Smith GC, Malone FD, Ball RH, Nyberg DA, Comstock CH, Hankins GD, Berkowitz RL, Gross SJ, Dugoff L, Craigo SD, Timor-Tritsch IE, Carr SR, Wolfe HM and D'Alton ME. Fetal growth in early pregnancy and risk of delivering low birth weight infant: prospective cohort study. *BMJ* 2007; **334**: 836.

Campbell S. An improved method of fetal cephalometry by ultrasound. *J Obstet Gynaecol Br Commonw* 1968; **75**: 568-576.

Campbell S. The prediction of fetal maturity by ultrasonic measurement of the biparietal diameter. *J Obstet Gynaecol Br Commonw* 1969; **76**: 603-609.

Campbell S and Thoms A. Ultrasound measurement of the fetal head to abdomen circumference ratio in the assessment of growth retardation. *Br J Obstet Gynaecol* 1977; **84**: 165-174.

Campbell S, Warsof SL, Little D and Cooper DJ. Routine ultrasound screening for the prediction of gestational age. *Obstet Gynecol* 1985; **65**: 613-620.

Campbell S. Early sonographic prenatal diagnosis. *Prenat Diagn* 2010; **30**: 613-615.

Caughey AB, Nicholson JM and Washington AE. First- vs second-trimester ultrasound: the effect on pregnancy dating and perinatal outcomes. *Am J Obstet Gynecol* 2008; **198: 703.e1–703.e6**.

Chalouhi GE, Bernard JP, Benoist G, Nasr B, Ville Y and Salomon LJ. A comparison of first trimester measurements for prediction of delivery date. *J Matern Fetal Neonatal Med* 2011; **24**: 51-57.

Chervenak FA, Skupski DW, Romero R, Myers MK, Smith-Levitin M, Rosenwaks Z and Thaler HT. How accurate is fetal biometry in the assessment of fetal age? *Am J Obstet Gynecol* 1998; **178**: 678-687.

Degani S. Fetal biometry: clinical, pathological, and technical considerations. *Obstet Gynecol Surv* 2001; **56**: 159-167.

Deter RL, Rossavik IK, Hill RM, Cortissoz C and Hadlock FP. Longitudinal studies of femur growth in normal fetuses. *J Clin Ultrasound* 1987; **15**: 299-305.

Dias T, Arcangeli T, Bhide A, Mahsud-Dornan S, Papageorghiou A and Thilaganathan B. Second-trimester assessment of gestational age in twins: validation of singleton biometry charts. *Ultrasound Obstet Gynecol* 2011; **37**: 34-37.

Drooger JC, Troe JW, Borsboom GJ, Hofman A, Mackenbach JP, Moll HA, Snijders RJ, Verhulst FC, Witteman JC, Steegers EA and Joung IM. Ethnic differences in prenatal growth and the association with maternal and fetal characteristics. *Ultrasound Obstet Gynecol* 2005; **26**: 115-122.

DuBose TJ and Hill LW. Physics and methodology of obstetrical sonography. In Fetal Sonography. DuBose TJ (ed). W.B. Saunders Company: Philadelphia, 1996; 73-75.

Dudley NJ and Chapman E. The importance of quality management in fetal measurement. *Ultrasound Obstet Gynecol* 2002; **19**: 190-196.

Eik-Nes SH, Grøttum P and Andersson NJ. Estimation of fetal weight by ultrasound measurement. II. Clinical application of a new formula. *Acta Obstet Gynecol Scand* 1982a; **61**: 307-312.

Eik-Nes SH, Grøttum P, Persson PH and Maršál K. Prediction of fetal growth deviation by ultrasonic biometry. I. Methodology. *Acta Obstet Gynecol Scand* 1982b; **61**: 53-58.

Eik-Nes SH and Grøttum P. Graviditetskalenderen Snurra. Scan-Med A/S, Drammen, Norway. 1983.

Eik-Nes SH, Økland O, Aure JC and Ulstein M. Ultrasound screening in pregnancy: a randomised controlled trial. *Lancet* 1984; **1**: 1347.

Eik-Nes SH, Salvesen KÅ, Økland O and Vatten LJ. Routine ultrasound fetal examination in pregnancy: the 'Ålesund' randomized controlled trial. *Ultrasound Obstet Gynecol* 2000; **15**: 473-478.

Eik-Nes SH, Grøttum P and Gjessing HK. Regarding "Term prediction with ultrasound: evaluation of a new dating curve for biparietal diameter". *Acta Obstet Gynecol Scand* 2006; **85**: 1276-1278; author reply 1278-1279.

Fellman V, Hellström-Westas L, Norman M, Westgren M, Källén K, Lagercrantz H, Maršál K, Serenius F and Wennergren M. One-year survival of extremely preterm infants after active perinatal care in Sweden. *Jama* 2009; **301**: 2225-2233.

Fox NS, Huang M and Chasen ST. Second-trimester fetal growth and the risk of poor obstetric and neonatal outcomes. *Ultrasound Obstet Gynecol* 2008; **32**: 61–65.

Gardosi J. Dating of pregnancy: time to forget the last menstrual period. *Ultrasound Obstet Gynecol* 1997; **9**: 367-368.

Gardosi J, Vanner T and Francis A. Gestational age and induction of labour for prolonged pregnancy. *Br J Obstet Gynaecol* 1997; **104**: 792-797.

Gardosi J and Geirsson RT. Routine ultrasound is the method of choice for dating pregnancy. *Br J Obstet Gynaecol* 1998; **105**: 933-936.

Geirsson RT. Ultrasound instead of last menstrual period as the basis of gestational age assignment. *Ultrasound Obstet Gynecol* 1991; **1**: 212-219.

Geirsson RT and Busby-Earle RM. Certain dates may not provide a reliable estimate of gestational age. *Br J Obstet Gynaecol* 1991; **98**: 108-109.

Gibson GB. Prolonged pregnancy. *British Medical Journal* 1955; **2**: 715-719.

Gjessing HK, Skjaerven R and Wilcox AJ. Errors in gestational age: evidence of bleeding early in pregnancy. *Am J Public Health* 1999; **89**: 213-218.

Gjessing HK and Grøttum P. Accuracy of second trimester fetal head circumference and biparietal diameter for predicting the time of spontaneous birth. *J Perinat Med* 2007; **35**: 350-351; author reply 351-352.

Gjessing HK, Grøttum P and 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.

Gjessing HK, Grøttum P, Økland I and Eik-Nes SH. Comment and reply on: prediction of the date of delivery based on first trimester ultrasound measurements: an independent method from estimated date of conception. *J Matern Fetal Neonatal Med* 2010; **23**: 944-946, author reply 946-947.

Goldenberg RL, Davis RO, Cutter GR, Hoffman HJ, Brumfield CG and Foster JM. Prematurity, postdates, and growth retardation: the influence of use of ultrasonography on reported gestational age. *American journal of obstetrics and gynecology* 1989; **160**: 462-470.

Goldstein RB, Filly RA and Simpson G. Pitfalls in femur length measurements. *J Ultrasound Med* 1987; **6**: 203-207.

Gottlieb AG and Galan HL. Nontraditional sonographic pearls in estimating gestational age. *Semin Perinatol* 2008; **32**: 154-160.

Grant A, Penn ZJ and Steer PJ. Elective or selective caesarean delivery of the small baby? A systematic review of the controlled trials. *Br J Obstet Gynaecol* 1996; **103**: 1197-1200.

Gülmezoglu AM, Crowther CA and Middleton P. Induction of labour for improving birth outcomes for women at or beyond term. *Cochrane Database Syst Rev* 2009. Issue 4. Art. No.: CD004945. DOI: 10.1002/14651858.CD004945.pub2.

Hadlock FP, Deter RL, Carpenter RJ and Park SK. Estimating fetal age: effect of head shape on BPD. *AJR Am J Roentgenol* 1981; **137**: 83-85.

Hadlock FP, Deter RL, Harrist RB and Park SK. Fetal biparietal diameter: rational choice of plane of section for sonographic measurement. *AJR Am J Roentgenol* 1982a; **138**: 871-874.

Hadlock FP, Deter RL, Harrist RB and Park SK. Fetal head circumference: relation to menstrual age. *AJR Am J Roentgenol* 1982b; **138**: 649-653.

Hadlock FP, Harrist RB, Deter RL and Park SK. Fetal femur length as a predictor of menstrual age: sonographically measured. *AJR Am J Roentgenol* 1982c; **138**: 875-878.

Hadlock FP, Deter RL, Harrist RB and Park SK. Estimating fetal age: computer-assisted analysis of multiple fetal growth parameters. *Radiology* 1984; **152**: 497-501.

Hall MH. Definitions used in relation to gestational age. *Paediatric and perinatal epidemiology* 1990; **4**: 123-128.

Harvey W. In Book *The works of William Harvey, M. D.* Translated from the Latin by R. Willis. The Sydenham Society (Digitized by Google): London, 1847; Page 523–544.

Heimstad R, Romundstad PR, Eik-Nes SH and Salvesen KA. Outcomes of pregnancy beyond 37 weeks of gestation. *Obstet Gynecol* 2006; **108**: 500-508.

Heimstad R, Skogvoll E, Mattsson LA, Johansen OJ, Eik-Nes SH and Salvesen KA. Induction of labor or serial antenatal fetal monitoring in postterm pregnancy: a randomized controlled trial. *Obstet Gynecol* 2007; **109**: 609-617.

Hilder L, Costeloe K and Thilaganathan B. Prolonged pregnancy: evaluating gestation-specific risks of fetal and infant mortality. *Br J Obstet Gynaecol* 1998; **105**: 169-173.

Hutchon DJ. "Back to future" for Hermaani Boerhaave, or, "A rational way to generate ultrasound scan charts for estimating the date of delivery". OBGYN.net: 1998; http://www.obgyn.net/us/cotm/9807/cotm_9807.htm [Accessed 15 March 2011]

Hutchon DJ and Ahmed F. Naegele's rule: a reappraisal. *Bjog* 2001; **108**: 775.

Høgberg U and Larsson N. Early dating by ultrasound and perinatal outcome. A cohort study. *Acta Obstet Gynecol Scand* 1997; **76**: 907-912.

Jackson RA, Gibson KA, Wu YW and Croughan MS. Perinatal outcomes in singletons following in vitro fertilization: a meta-analysis. *Obstet Gynecol* 2004; **103**: 551-563.

Jago JR, Whittingham TA and Heslop R. The influence of ultrasound scanner beam width on femur length measurements. *Ultrasound Med Biol* 1994; **20**: 699-703.

Johnsen SL, Rasmussen S, Sollien R and Kiserud T. Fetal age assessment based on ultrasound head biometry and the effect of maternal and fetal factors. *Acta Obstet Gynecol Scand* 2004; **83**: 716-723.

Johnsen SL, Rasmussen S, Sollien R and Kiserud T. Fetal age assessment based on femur length at 10-25 weeks of gestation, and reference ranges for femur length to head circumference ratios. *Acta Obstet Gynecol Scand* 2005; **84**: 725-733.

Kagan KO, Etchegaray A, Zhou Y, Wright D and Nicolaides KH. Prospective validation of first-trimester combined screening for trisomy 21. *Ultrasound Obstet Gynecol* 2009; **34**: 14-18.

Kieler H, Axelsson O, Nilsson S and Waldenstrom U. The length of human pregnancy as calculated by ultrasonographic measurement of the fetal biparietal diameter. *Ultrasound Obstet Gynecol* 1995; **6**: 353-357.

Kirkegaard I, Henriksen TB and Uldbjerg N. Early fetal growth, PAPP-A and free beta-hCG in relation to risk of delivering a small-for-gestational age infant. *Ultrasound Obstet Gynecol* 2011; **37**: 341-347.

Kiserud T and Rasmussen S. [Assessment of gestational age using ultrasound – can the method be improved?]. *Tidsskr Nor Laegeforen* 1999; **119**: 4331-4334.

Koster MP, Van Leeuwen-Spruijt M, Wortelboer EJ, Stoutenbeek P, Elvers LH, Loeber JG, Visser GH and Schielen PC. Lack of standardization in determining gestational age for prenatal screening. *Ultrasound Obstet Gynecol* 2008; **32**: 607-611.

Kurmanavicius J, Wright EM, Royston P, Zimmermann R, Huch R, Huch A and Wisser J. Fetal ultrasound biometry: 2. Abdomen and femur length reference values. *Br J Obstet Gynaecol* 1999; **106**: 136-143.

Kurtz A and Goldberg B. Fetal Head Measurements. In Book: Obstetrical Measurements in Ultrasound. A Reference Manual. Chicago; Year Book Medical, 1988.

Kvande L. *Bilete av svangerskap – bilete av foster*. In Thesis*;* Norwegian University of Science and Technology: Trondheim, 2008.

Källén K. Mid-trimester ultrasound prediction of gestational age: advantages and systematic errors. *Ultrasound Obstet Gynecol* 2002; **20**: 558-563.

Källén K. Increased risk of perinatal/neonatal death in infants who were smaller than expected at ultrasound fetometry in early pregnancy. *Ultrasound Obstet Gynecol* 2004; **24**: 30-34.

Larsen T, Nguyen TH, Greisen G, Engholm G and Møller H. Does a discrepancy between gestational age determined by biparietal diameter and last menstrual period sometimes signify early intrauterine growth retardation? *Bjog* 2000; **107**: 238-244.

Longo D, DeFigueiredo D, Cicero S, Sacchini C and Nicolaides KH. Femur and humerus length in trisomy 21 fetuses at 11-14 weeks of gestation. *Ultrasound Obstet Gynecol* 2004; **23**: 143-147.

Loughna P, Chitty L, Evans T and Chudleigh T. Fetal size and dating: charts recommended for clinical obstetric practice. *Ultrasound* 2009; **17**: 161-167.

Lynch CD and Zhang J. The research implications of the selection of a gestational age estimation method. *Paediatr Perinat Epidemiol* 2007; **21 Suppl 2**: 86-96.

Mandruzzato G, Alfirevic Z, Chervenak F, Gruenebaum A, Heimstad R, Heinonen S, Levene M, Salvesen K, Saugstad O, Skupski D and Thilaganathan B. Guidelines for the management of postterm pregnancy. *J Perinat Med* 2010; **38**: 111-119.

Markestad T, Kaaresen PI, Rønnestad A, Reigstad H, Lossius K, Medbø S, Zanussi G, Engelund IE, Skjaerven R and Irgens LM. Early death, morbidity, and need of treatment among extremely premature infants. *Pediatrics* 2005; **115**: 1289-1298.

Maršál K. Obstetric management of intrauterine growth restriction. *Best Pract Res Clin Obstet Gynaecol* 2009; **23**: 857-870.

Miljeteig I, Markestad T and Norheim OF. Physicians' use of guidelines and attitudes to withholding and withdrawing treatment for extremely premature neonates in Norway. *Acta Paediatr* 2007; **96**: 825-829.

Mittendorf R, Williams MA, Berkey CS and Cotter PF. The length of uncomplicated human gestation. *Obstet Gynecol* 1990; **75**: 929-932.

Mongelli M, Wilcox M and Gardosi J. Estimating the date of confinement: ultrasonographic biometry versus certain menstrual dates. *Am J Obstet Gynecol* 1996; **174**: 278-281.

Mongelli M, Yuxin NG, Biswas A and Chew S. Accuracy of ultrasound dating formulae in the late second-trimester in pregnancies conceived with in-vitro fertilization. *Acta Radiol* 2003; **44**: 452-455.

Mongelli M, Chew S, Yuxin NG and Biswas A. Third-trimester ultrasound dating algorithms derived from pregnancies conceived with artificial reproductive techniques. *Ultrasound Obstet Gynecol* 2005; **26**: 129-131.

Morin I, Morin L, Zhang X, Platt RW, Blondel B, Breart G, Usher R and Kramer MS. Determinants and consequences of discrepancies in menstrual and ultrasonographic gestational age estimates. *Bjog* 2005; **112**: 145-152.

Mul T, Mongelli M and Gardosi J. A comparative analysis of second-trimester ultrasound dating formulae in pregnancies conceived with artificial reproductive techniques. *Ultrasound Obstet Gynecol* 1996; **8**: 397-402.

National Collaborating Centre for Women's and Children's Health: Antenatal care. Routine care for the healthy pregnant woman. NICE clinical guideline 62. National Institute for Health and Clinical Excellence: London, 2010; http://www.nice.org.uk/nicemedia/live/11947/40115/40115.pdf [Accessed 10 March 2011]

Nguyen TH, Larsen T, Engholm G and Møller H. Evaluation of ultrasound-estimated date of delivery in 17,450 spontaneous singleton births: do we need to modify Naegele's rule? *Ultrasound Obstet Gynecol* 1999; **14**: 23-28.

Nguyen TH, Larsen T, Engholm G and Møller H. A discrepancy between gestational age estimated by last menstrual period and biparietal diameter may indicate an increased risk of fetal death and adverse pregnancy outcome. *Bjog* 2000; **107**: 1122-1129.

Nicolaides KH. Screening for fetal aneuploidies at 11 to 13 weeks. *Prenatal diagnosis* 2011a; **31**: 7-15.

Nicolaides KH. A model for a new pyramid of prenatal care based on the 11 to 13 weeks' assessment. *Prenatal diagnosis* 2011b; **31**: 3-6.

Nicolaides KH. Turning the pyramid of prenatal care. *Fetal Diagn Ther* 2011c; **29**: 183-196.

Norwitz ER, Snegovskikh VV and Caughey AB. Prolonged pregnancy: when should we intervene? *Clin Obstet Gynecol* 2007; **50**: 547-557.

Norwitz ER. Postterm pregnancy. UpToDate®: 2011; http://www.uptodate.com/contents/postterm-pregnancy?source=preview&selectedTitle=1%7E27&anchor=H1#H1 [Accessed 25 May 2011]

Nyberg DA. May all your femurs be long! *Ultrasound Obstet Gynecol* 2008; **31**: 489-492.

Nägele FC. In Book *Erfahrungen und Abhandlungen aus dem Gebiethe der Krankheiten des Weiblichen Geschlechtes. Nebst Grundzügen einer Methodenlehre der Geburtshülfe*. T. Loeffler: Mannheim, 1812; Page 280–281.

Nägele FC. In Book *Lehrbuch der Geburtshülfe für Hebammen*. Akademische Verlagsbuchhandlung von J. C. B. Mohr: Heidelberg, 1868; Page 89–90.

Offerdal K, Jebens N, Syvertsen T, Blaas H-GK, Johansen OJ and Eik-Nes SH. Prenatal ultrasound detection of facial clefts: a prospective study of 49,314 deliveries in a non-selected population in Norway. *Ultrasound Obstet Gynecol* 2008; **31**: 639-646.

Olsen O and Aaroe Clausen J. Routine ultrasound dating has not been shown to be more accurate than the calendar method. *Br J Obstet Gynaecol* 1997; **104**: 1221-1222.

Ott WJ. Accurate gestational dating. *Obstet Gynecol* 1985; **66**: 311-315.

Pedersen NG, Figueras F, Wojdemann KR, Tabor A and Gardosi J. Early fetal size and growth as predictors of adverse outcome. *Obstet Gynecol* 2008a; **112**: 765-771.

Pedersen NG, Wojdemann KR, Scheike T and Tabor A. Fetal growth between the first and second trimesters and the risk of adverse pregnancy outcome. *Ultrasound Obstet Gynecol* 2008b; **32**: 147-154.

Perloff WH and Steinberger E. In Vivo Survival of Spermatozoa in Cervical Mucus. *Am J Obstet Gynecol* 1964; **88**: 439-442.

Persson PH and Kullander S. Long-term experience of general ultrasound screening in pregnancy. *Am J Obstet Gynecol* 1983; **146**: 942-947.

Persson PH. Ultrasound dating of pregnancy – still controversial? *Ultrasound Obstet Gynecol* 1999; **14**: 9-11.

Pexsters A, Daemen A, Bottomley C, Van Schoubroeck D, De Catte L, De Moor B, D'Hooghe T, Lees C, Timmerman D and Bourne T. New crown-rump length curve based on over 3500 pregnancies. *Ultrasound Obstet Gynecol* 2010; **35**: 650-655.

Pignotti MS. Extremely preterm births: recommendations for treatment in European countries. *Arch Dis Child Fetal Neonatal Ed* 2008; **93**: 403-406.

Pretorius DH, Nelson TR and Manco-Johnson ML. Fetal age estimation by ultrasound: the impact of measurement errors. *Radiology* 1984; **152**: 763-766.

R. Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria, 2010; http://www.R-project.org. [Accessed 15 January 2011]

Rijken MJ, Lee SJ, Boel ME, Papageorghiou AT, Visser GH, Dwell SL, Kennedy SH, Singhasivanon P, White NJ, Nosten F and McGready R. Obstetric ultrasound scanning by local health workers in a refugee camp on the Thai-Burmese border. *Ultrasound Obstet Gynecol* 2009; **34**: 395-403.

Robinson HP. Sonar measurement of fetal crown-rump length as means of assessing maturity in first trimester of pregnancy. *Br Med J* 1973; **4**: 28-31.

Robinson HP and Fleming JE. A critical evaluation of sonar "crown-rump length" measurements. *Br J Obstet Gynaecol* 1975; **82**: 702-710.

Romundstad LB, Romundstad PR, Sunde A, von During V, Skjaerven R, Gunnell D and Vatten LJ. Effects of technology or maternal factors on perinatal outcome after assisted fertilisation: a population-based cohort study. *Lancet* 2008; **372**: 737-743.

Roos N, Sahlin L, Ekman-Ordeberg G, Kieler H and Stephansson O. Maternal risk factors for postterm pregnancy and cesarean delivery following labor induction. *Acta Obstet Gynecol Scand* 2010; **89**: 1003-1010.

Rosati P, Guariglia L and Capelli G. A new mathematical formula for predicting long bone length in early pregnancy. *Ultrasound Obstet Gynecol* 2002; **19**: 184-189.

Rossavik IK and Fishburne JI. Conceptional age, menstrual age, and ultrasound age: a second-trimester comparison of pregnancies of known conception date with pregnancies dated from the last menstrual period. *Obstet Gynecol* 1989; **73**: 243-249.

Rossen J, Økland I, Nilsen OB and Eggebø TM. Is there an increase of postpartum hemorrhage, and is severe hemorrhage associated with more frequent use of obstetric interventions? *Acta Obstet Gynecol Scand* 2010; **89**: 1248-1255.

Royston P and Wright EM. How to construct 'normal ranges' for fetal variables. *Ultrasound Obstet Gynecol* 1998; **11**: 30-38.

Sahota DS, Kagan KO, Lau TK, Leung TY and Nicolaides KH. Customized birth weight: coefficients and validation of models in a UK population. *Ultrasound Obstet Gynecol* 2008; **32**: 884-889.

Salomon LJ, Bernard JP, Duyme M, Buvat I and Ville Y. The impact of choice of reference charts and equations on the assessment of fetal biometry. *Ultrasound Obstet Gynecol* 2005; **25**: 559-565.

Salomon LJ, Bernard JP, Duyme M, Doris B, Mas N and Ville Y. Feasibility and reproducibility of an image-scoring method for quality control of fetal biometry in the second trimester. *Ultrasound Obstet Gynecol* 2006a; **27**: 34-40.

Salomon LJ, Duyme M, Crequat J, Brodaty G, Talmant C, Fries N and Althuser M. French fetal biometry: reference equations and comparison with other charts. *Ultrasound Obstet Gynecol* 2006b; **28**: 193-198.

Salomon LJ, Winer N, Bernard JP and Ville Y. A score-based method for quality control of fetal images at routine second-trimester ultrasound examination. *Prenat Diagn* 2008; **28**: 822-827.

Salomon LJ, Pizzi C, Gasparrini A, Bernard JP and Ville Y. Prediction of the date of delivery based on first trimester ultrasound measurements: An independent method from estimated date of conception. *J Matern Fetal Neonatal Med* 2010; **23**: 1-9.

Salomon LJ, Alfirevic Z, Berghella V, Bilardo C, Hernandez-Andrade E, Johnsen SL, Kalache K, Leung KY, Malinger G, Munoz H, Prefumo F, Toi A and Lee W. Practice guidelines for performance of the routine mid-trimester fetal ultrasound scan. *Ultrasound Obstet Gynecol* 2011; **37**: 116-126.

Saltvedt S, Almstrom H, Kublickas M, Reilly M, Valentin L and Grunewald C. Ultrasound dating at 12-14 or 15-20 weeks of gestation? A prospective cross-validation of established dating formulae in a population of in-vitro fertilized pregnancies randomized to early or late dating scan. *Ultrasound Obstet Gynecol* 2004; **24**: 42-50.

Saugstad OD. When newborn infants are bound to die. *Acta Paediatr* 2005; **94**: 1535-1537.

Savitz DA, Terry JW, Jr., Dole N, Thorp JM, Jr., Siega-Riz AM and Herring AH. Comparison of pregnancy dating by last menstrual period, ultrasound scanning, and their combination. *Am J Obstet Gynecol* 2002; **187**: 1660-1666.

Sladkevicius P, Saltvedt S, Almström H, Kublickas M, Grunewald C and Valentin L. Ultrasound dating at 12-14 weeks of gestation. A prospective cross-validation of established dating formulae in in-vitro fertilized pregnancies. *Ultrasound Obstet Gynecol* 2005; **26**: 504-511.

Snijders RJ and Nicolaides KH. Fetal biometry at 14-40 weeks' gestation. *Ultrasound Obstet Gynecol* 1994; **4**: 34-48.

Snijders RJ, Noble P, Sebire N, Souka A and Nicolaides KH. UK multicentre project on assessment of risk of trisomy 21 by maternal age and fetal nuchal-translucency thickness at 10-14 weeks of gestation. Fetal Medicine Foundation First Trimester Screening Group. *Lancet* 1998; **352**: 343-346.

Taipale P and Hiilesmaa V. Predicting delivery date by ultrasound and last menstrual period in early gestation. *Obstet Gynecol* 2001; **97**: 189-194.

Tegnander E and Eik-Nes SH. The examiner's ultrasound experience has a significant impact on the detection rate of congenital heart defects at the second-trimester fetal examination. *Ultrasound Obstet Gynecol* 2006; **28**: 8-14.

The Fetal Medicine Foundation: Online Education. Pyramid of care. London, 2010a; http://www.fetalmedicine.com/fmf/online-education/08-pyramid-of-care/ [Accessed 15 May 2011]

The Fetal Medicine Foundation: Training & Certification. Certificates of competence. London, 2010b; http://www.fetalmedicine.com/fmf/training-certification/certificates-of-competence/ [Accessed 12 May 2011]

Thorsell M, Kaijser M, Almstrom H and Andolf E. Expected day of delivery from ultrasound dating versus last menstrual period--obstetric outcome when dates mismatch. *Bjog* 2008; **115**: 585-589.

Tunón K, Eik-Nes SH and Grøttum P. A comparison between ultrasound and a reliable last menstrual period as predictors of the day of delivery in 15 000 examinations. *Ultrasound Obstet Gynecol* 1996; **8**: 178-185.

Tunón K, Eik-Nes SH and Grøttum P. The impact of fetal, maternal and external factors on prediction of the day of delivery by the use of ultrasound. *Ultrasound Obstet Gynecol* 1998; **11**: 99-103.

Tunón K, Eik-Nes SH and Grøttum P. Fetal outcome when the ultrasound estimate of the day of delivery is more than 14 days later than the last menstrual period estimate. *Ultrasound Obstet Gynecol* 1999a; **14**: 17-22.

Tunón K, Eik-Nes SH and Grøttum P. Fetal outcome in pregnancies defined as post-term according to the last menstrual period estimate, but not according to the ultrasound estimate. *Ultrasound Obstet Gynecol* 1999b; **14**: 12-16.

Tunón K, Eik-Nes SH, Grøttum P, Von Düring V and Kahn JA. Gestational age in pregnancies conceived after in vitro fertilization: a comparison between age assessed from oocyte retrieval, crown-rump length and biparietal diameter. *Ultrasound Obstet Gynecol* 2000; **15**: 41-46.

Verburg BO, Mulder PG, Hofman A, Jaddoe VW, Witteman JC and Steegers EA. Intra- and interobserver reproducibility study of early fetal growth parameters. *Prenatal diagnosis* 2008a; **28**: 323-331.

Verburg BO, Steegers EA, De Ridder M, Snijders RJ, Smith E, Hofman A, Moll HA, Jaddoe VW and Witteman JC. New charts for ultrasound dating of pregnancy and assessment of fetal growth: longitudinal data from a population-based cohort study. *Ultrasound Obstet Gynecol* 2008b; **31**: 388-396.

Ville Y. 'Ceci n'est pas une Èchographie': a plea for quality assessment in prenatal ultrasound. *Ultrasound Obstet Gynecol* 2008; **31**: 1-5.

Wennerholm UB, Hagberg H, Brorsson B and Bergh C. Induction of labor versus expectant management for post-date pregnancy: is there sufficient evidence for a change in clinical practice? *Acta Obstet Gynecol Scand* 2009; **88**: 6-17.

Whitworth M, Bricker L, Neilson JP and Dowswell T. Ultrasound for fetal assessment in early pregnancy. *Cochrane Database Syst Rev* 2010. Issue 4. Art. No.: CD007058. DOI: 10.1002/14651858.CD007058.pub2.

Wilcox AJ, Weinberg CR and Baird DD. Timing of sexual intercourse in relation to ovulation. Effects on the probability of conception, survival of the pregnancy, and sex of the baby. *N Engl J Med* 1995; **333**: 1517-1521.

Yang H, Kramer MS, Platt RW, Blondel B, Breart G, Morin I, Wilkins R and Usher R. How does early ultrasound scan estimation of gestational age lead to higher rates of preterm birth? *Am J Obstet Gynecol* 2002; **186**: 433-437.

Zeitlin J, Blondel B, Alexander S and Breart G. Variation in rates of postterm birth in Europe: reality or artefact? *Bjog* 2007; **114**: 1097-1103.

Økland I, Gjessing HK, Grøttum P and Eik-Nes SH. Biases of traditional term prediction models: results from different sample-based models evaluated on 41 343 ultrasound examinations. *Ultrasound Obstet Gynecol* 2010; **36**: 728-734.

Økland I, Bjåstad TG, Johansen TF, Gjessing HK, Grøttum P and Eik-Nes SH. Narrowed beam width in newer ultrasound machines shortens measurements in the lateral direction: fetal measurement charts may be obsolete. *Ultrasound Obstet Gynecol* 2011a; **38**: 82-87.

Økland I, Gjessing HK, Grøttum P, Eggebø TM and Eik-Nes SH. A new population-based term prediction model vs. two traditional sample-based models:

validation on 9046 ultrasound examinations. *Ultrasound Obstet Gynecol* 2011b; **37**: 207-213.

Økland I, Nakling J, Gjessing HK, Grøttum P and Eik-Nes SH. Advantages of the population-based approach to pregnancy dating demonstrated with results from 23 020 ultrasound examinations. *Ultrasound Obstet Gynecol* 2011c: DOI: 10.1002/uog.10081.

Paper I

Ultrasound Obstet Gynecol 2010; **36**: 728–734
Published online 15 November 2010 in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/uog.7707

Biases of traditional term prediction models: results from different sample-based models evaluated on 41 343 ultrasound examinations

I. ØKLAND*†, H. K. GJESSING‡§, P. GRØTTUM¶ and S. H. EIK-NES*#

**National Center for Fetal Medicine, Women and Children's Center, St Olavs University Hospital and #Norwegian University of Science and Technology, Trondheim, †Department of Obstetrics and Gynecology, Stavanger University Hospital, Stavanger, ‡Division of Epidemiology, Norwegian Institute of Public Health and ¶Section of Medical Informatics, University of Oslo, Oslo and §Department of Public Health and Primary Health Care, University of Bergen, Bergen, Norway*

KEYWORDS: date of delivery; fetal measurements; gestational age; population-based; routine ultrasound examination; second trimester; term prediction

ABSTRACT

Objective *To evaluate two Norwegian traditional, sample-based term prediction models as applied to the data from a large population-based registry. The two models were also compared with an established German model.*

Methods *Our database included information from 41 343 non-selected ultrasound scans registered over the years 1987–2005. The prediction models were applied to measurements from the ultrasound examinations, and the resulting term predictions were compared with the actual times of the deliveries. The median bias (the difference between the true and the predicted date of delivery) was calculated for each model, both for the study population as a whole and for subgroups of measurements of biparietal diameter (BPD) and femur length (FL). Secondary measures, i.e. proportion of births within ± 14 days and the rates of preterm and post-term deliveries, were also assessed.*

Results *The analyses showed that the models had significant biases, predicting delivery date either too late or too early. For each model the size of the bias varied, depending on the fetal size at the time of the examination; the extremes were minus 4 and plus 4 days for the BPD-based predictions. There were similar results with the FL-based predictions.*

Conclusion *Term predictions made with traditional sample-based models had significant biases that varied over each method's measurement range. These models have important shortcomings, probably because of strict selection criteria in the process of constructing the models, and because the methods primarily aim at estimating the last menstrual period-based day of conception, not the day of birth. Copyright © 2010 ISUOG. Published by John Wiley & Sons, Ltd.*

INTRODUCTION

Correct fetal age assessment is an important part of modern pregnancy care. The monitoring and management of every pregnancy that does not progress normally require detailed knowledge of the length of the pregnancy. To give optimal care to a baby born preterm, the midwife, obstetrician and pediatrician need to know the true gestational age[1]. Closer to term and in the post-term period an understanding of the risk estimates for the fetus, related to its predicted term, is fundamental[2]. Thus, information about the fetal age affects management schemes throughout pregnancy, as well as the organization surrounding the delivery. Also for social reasons, the knowledge of the expected day of delivery is important to the pregnant woman.

Ultrasound measurements of selected fetal parameters for the estimation of gestational age were introduced around 1970[3] and are now the method of choice for dating a pregnancy[4–6]. However, the traditional sample-based ultrasound models for predicting date of delivery have been shown to vary in predictive quality[7,8], as has also been shown for one of the models in this study[9]. Conventional dating charts are based mainly on measurements from a carefully selected population

Correspondence to: I. Økland, Department of Obstetrics and Gynecology, Stavanger University Hospital, PO Box 8100, N-4068 Stavanger, Norway (e-mail: iok@lyse.net)

Accepted: 26 May 2010

of pregnant women, constituting a limited reference sample. In addition, the models estimate a date of the last menstrual period (LMP) from the fetal ultrasound measurements; the subsequent prediction of day of delivery is done simply by adding approximately 280 days to the estimated, and therefore 'artificial', LMP date[10]. The focus on determining the correct date of the LMP significantly increases the risk of an inaccurate term prediction. These factors are intrinsic weaknesses in the design of sample-based models. Consequently, a reappraisal of the traditional dating models is called for.

The aim of this study was to evaluate three sample-based models for prediction of date of delivery on data from the large population-based registry of approximately 40 000 pregnancies at the National Center for Fetal Medicine (NCFM) in Norway.

SUBJECTS AND METHODS

Subjects

The women included in this study had their fetal ultrasound examinations at the NCFM at St Olavs University Hospital in Trondheim, Norway, and subsequently were delivered at the hospital. The study population was non-selected, coming from a geographically well-defined area consisting of the city of Trondheim and eight surrounding municipalities. Within this population, approximately 97% of the pregnant women were examined and delivered in Trondheim. Data were collected over the period 1987–2005 and included a total of 50 533 second-trimester examinations of 45 343 pregnancies, each with complete information about the date of the ultrasound scan and the date of delivery, and with a fetal biparietal diameter (BPD) in the range 25–60 mm or a femur length (FL) in the range 11–42 mm at the time of the ultrasound examination. In accordance with the Norwegian practice of routine ultrasound examinations, the majority of the examinations were from pregnancy weeks 17–19.

Complicated pregnancies related to stillbirths ($n = 478$), diagnosed anomalies ($n = 1935$), multiple pregnancies ($n = 696$) and induction of labor for reasons other than post-term pregnancy ($n = 4944$) were excluded. As a conservative measure and irrespective of the result of the examination, we additionally excluded 1137 scans performed for various indications that could possibly be related to abnormal fetal growth, e.g. maternal concern about the growth of the fetus. Thus, 41 343 examinations in 36 982 pregnancies remained for analysis. Women induced for post-term pregnancy ($n = 772$) were not excluded. The median age of the women was 28 years, 44.1% were primiparous and 21.6% reported daily smoking.

Prediction models

Three different models for predicting date of delivery – two Norwegian and one German – were evaluated in this study. All three models were constructed from limited reference samples, by regressing the LMP-based gestational age on the measured fetal size. In clinical use, the date of conception is predicted from the fetal size, and the estimated date of delivery (EDD) is then predicted by adding an assumed standard length of pregnancy.

The obstetric wheel 'Snurra' (Trondheim–1984)[11] was the only dating method in use in Norway from 1984 until recently. It predicts term from BPD measurements between 38 and 60 mm only. The model was constructed from a population of 90 women with anticipated normal pregnancies, all of them carefully selected regarding menstrual history. The women were included in a prospective, longitudinal study, and measurements were taken from each fetus approximately 10 times. Fourth-order polynomial regression analysis was used to establish the curves.

A more recently developed obstetric wheel 'Terminhjulet' (Bergen–2004)[12,13] is based on the same principles as Trondheim–1984[14], but uses a newer data sample and fractional polynomial regression analysis. This model was constructed from a prospective, cross-sectional study of 650 healthy women with regular menstrual periods and singleton, uncomplicated pregnancies. From Bergen–2004 the date of delivery may be predicted from BPD (14–60 mm), FL (2–44 mm) or head circumference (50–134 mm) measurements.

The third model, constructed by Hansmann in 1976[15] (Hansmann–1976), applied longitudinal data from 1348 selected singleton pregnancies with reliable LMP data to construct the BPD-based dating model and a term prediction wheel. Hansmann–1976 used a traditional regression model.

The BPD prediction tables for Trondheim–1984 and Hansmann–1976 are available from the respective prediction models[11,15]. For Bergen–2004, the BPD and FL tables in the articles show the computed gestational age, but the accompanying published formulae[12,13] are in error. The correct formulae, which we used in this study, were:

$$\log GA = 2.507 - 1.333 \times (BPD)^{-1/2} + 0.01393 \times BPD$$

$$GA = 8.625 + 1.395 \times (FL)^{1/2} + 0.003684 \times (FL)^2,$$

where GA is gestational age in weeks. (S. L. Johnsen, personal communication.)

For both Trondheim–1984 and Bergen–2004 the predicted date of delivery is found by adding 282 days to the deduced LMP date, while the Hansmann–1976 model assumes a pregnancy length of 280 days.

Ultrasound examinations

More than 30 experienced and formally trained midwives performed the fetal ultrasound examinations, and the data were prospectively registered in an electronic database.

In general, all clinical problems were primarily managed according to the Trondheim–1984 prediction model; post-term pregnant women were scheduled for induction

Ultrasound Obstet Gynecol 2010; **36**: 728–734.

of labor after 13 days past EDD (corresponding to > 295 days), and induction for clinical problems possibly related to post-term pregnancy was not acknowledged as such until 12 days past EDD (≥ 294 days).

The BPD was measured as described by Campbell and Thoms[16], through a horizontal section of the fetal head at the level of the cavum septi pellucidi, from outer to outer contours of the parietal bones. The mean of three BPD measurements was used for the calculation. The FL was measured with the femoral diaphysis in a longitudinal section, in accordance with the method described by Goldstein *et al.*[17] and Hadlock *et al.*[18]. The femur was measured three times and the longest measurement was used[18–20]. Both BPD and FL measurements were rounded to the nearest millimeter.

The Bergen–2004 model used the mean of three FL measurements rather than adhering to the international practice of using the longest of three FL measurements[18–20]. To address this problem we carried out a separate analysis to explore what impact the Bergen–2004 measurement technique had on the FL-based term predictions.

The following ultrasound scanners were used: Hitachi EUB–410, EUB–415, EUB–6000 and EUB–6500 (Tokyo, Japan), Vingmed System Five (Vingmed Sound, Horten, Norway) and Logic 500 (GE Healthcare, Milwaukee, WI, USA) with curvilinear 3.5–5-MHz transducers. The sound velocity was calibrated to 1540 m/s.

Statistical methods

The prediction models were applied to measurements from the ultrasound examinations. The 38 266 measurements with a BPD in the range 38–60 mm were used to predict the date of delivery according to Trondheim–1984, and the 40 248 measurements with a BPD in the range 25–60 mm were used with Bergen–2004 and with Hansmann–1976. The 39 989 accessible FL measurements from 11 to 42 mm were used with Bergen–2004 and the resulting predictions were compared with the actual times of the succeeding deliveries.

A correction for the time period[21] was included for Bergen–2004, as this model was developed in recent years with modern ultrasound machines and there is reason to believe that the narrower beam width of newer ultrasound devices might influence some measurements[22].

Our primary measure for assessing prediction quality was the median bias, defined as the difference between the true and the predicted date of delivery, thus reflecting the systematic error in prediction. A positive bias means that the birth took place later than predicted; a negative bias means it took place before predicted. A systematic positive bias will lead to an inflated number of births defined as post-term and, similarly, a negative bias will lead to an apparent decrease in the post-term group. To measure this effect of misclassification, we also computed secondary measures, i.e. the proportion of births within ± 14 days of the EDD, and the rates of preterm (24 days or more before EDD) and post-term (14 days or more after EDD) deliveries. For the Hansmann–1976 model, only the median bias was analyzed.

Bias may vary over the range of inclusion, and a variable bias, depending on the fetal size at the time of the ultrasound examination, may be overlooked if only the overall bias is calculated[23]. Therefore, for each method, median bias and secondary measures were calculated for the study population as a whole and for three subgroups of BPD/FL measurements.

Confidence intervals for the bias values were computed using bootstrapping with 2000 replications. All analyses and graphics were produced in the R statistical programming environment[24].

RESULTS

BPD-based predictions

Table 1 shows the median biases and the values of the secondary measures when the date of delivery was predicted with each of the two Norwegian models for the study group as a whole, and for different groups with BPD ranges corresponding to a gestational age below 18 weeks (38–43 mm), 18 weeks (44–46 mm) and above 18 weeks (47–60 mm), respectively. For Bergen–2004 we also analyzed the results for the BPD range 25–37 mm.

Figure 1 shows the birth distribution histograms relative to the term predicted by the two Norwegian sample-based models, with the optimized distribution curve of birth as computed from the population-based model[21] (superimposed line). Figure 2 shows the analyses of the median biases for all three models. The median biases of the BPD-based predictions varied between minus 4 and plus 4 days. Within each model, there were substantial variations in median biases over the inclusion range; Hansmann–1976 varied between −2 and 4 days, Trondheim–1984 between −4.5 and 1 day and Bergen–2004 between −0.5 and 4 days. Hansmann–1976 showed biases very similar to Bergen–2004 in the main BPD-prediction span.

FL-based predictions

The Bergen–2004 model was applied to the FL measurements. The birth distribution histogram relative to the date of delivery predicted by this model, with the optimized birth distribution curve as computed from the population-based model[21] as the superimposed line, is shown in Figure 3.

In Table 2 the median biases, the proportion of births within ± 14 days and rates of pre- and post-term deliveries are presented for four different groups with FL ranges corresponding to a gestational age of less than 16 weeks (11–22 mm), 16–18 weeks (23–26 mm), 18 weeks (27–29 mm) and greater than 18 weeks (30–42 mm), respectively, and for the study group as a whole.

The median bias for Bergen–2004 increased with increasing FL (Figure 4). Even using the same model

Table 1 Prediction of date of delivery from biparietal diameter (BPD) with the two Norwegian models: median bias, proportion of births within ± 14 days, and proportions of preterm and post-term deliveries for different BPD-range groups and for the study group as a whole

			Primary measure	*Secondary measures*		
BPD (mm)	n	*Method*	*Median bias (95% CI) (days)*	*± 14 days (%)*	*Preterm (%)*	*Post-term (%)*
25–37	1982	Bergen–2004	1.70 (1.34 to 2.26)	85.9	3.5	6.8
38–43	10 734	Trondheim–1984	−3.07 (−3.29 to −2.87)	85.6	4.7	2.3
		Bergen–2004	0.17 (0.15 to 0.22)	87.1	3.3	5.1
44–46	13 821	Trondheim–1984	−1.28 (−1.46 to −1.12)	87.7	4.0	3.5
		Bergen–2004	1.40 (1.40 to 1.52)	86.9	3.1	6.7
47–60	13 711	Trondheim–1984	−0.16 (−0.34 to 0.03)	87.0	3.7	4.9
		Bergen–2004	3.07 (2.86 to 3.31)	84.4	2.7	10.0
38–60	38 266	Trondheim–1984	−1.43 (−1.54 to −1.33)	86.9	4.1	3.7
		Bergen–2004	1.52 (1.40 to 1.68)	86.1	3.0	7.4

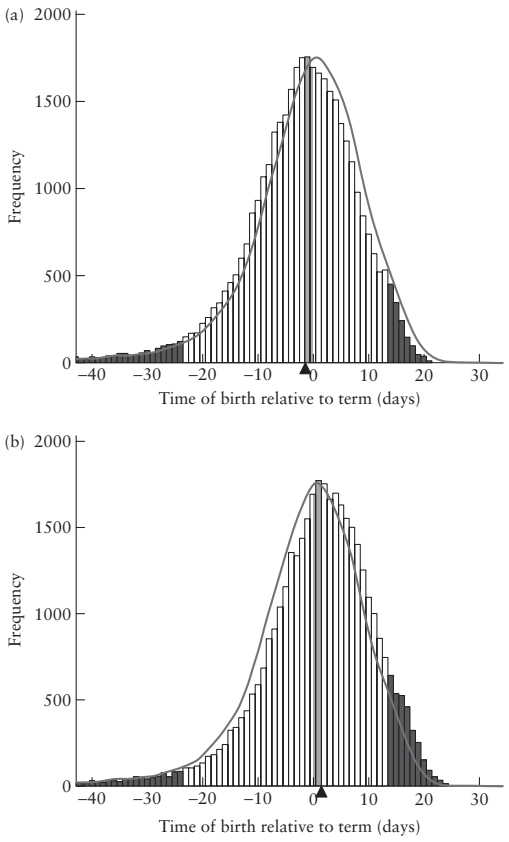


Figure 1 Birth distribution histograms showing the median time of birth (▲) relative to the predicted term (0) from 38 266 measurements of biparietal diameter in the range of 38–60 mm for (a) the Trondheim–1984 model and (b) the Bergen–2004 model, with the optimized birth distribution curve superimposed (continuous line). □, Days of median bias; ■, preterm and post-term deliveries.

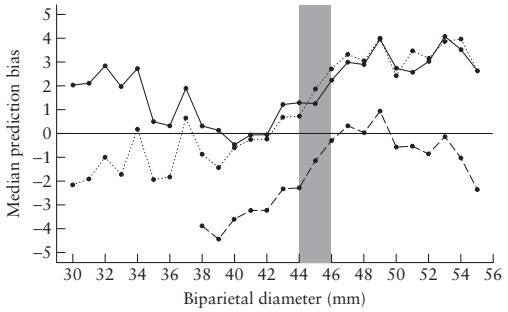


Figure 2 Median biases for the three models related to different measurements of biparietal diameter (BPD): Trondheim–1984 (– – –); Bergen–2004 (——); Hansmann–1976 (.....). □, Central BPD area (week 18).

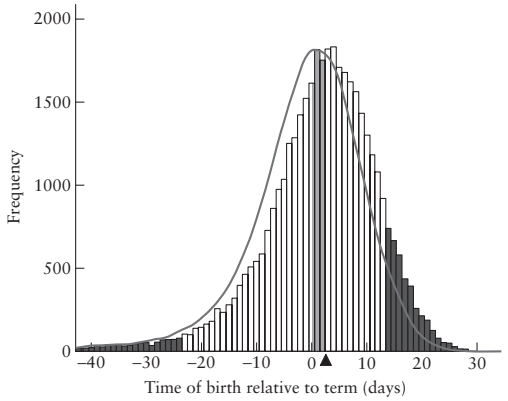


Figure 3 Birth distribution histogram showing the median time of birth (▲) relative to the predicted term (0) from 39 989 measurements of femur length for the Bergen–2004 model, with the optimized birth distribution curve superimposed (continuous line). □, Days of median bias; ■, preterm and post-term deliveries.

Table 2 Prediction of date of delivery from femur length (FL) with the Bergen–2004 model: median bias, proportion of births within ± 14 days, and proportions of preterm and post-term deliveries for four different FL-range groups and for the study group as a whole

		Primary measure	*Secondary measures*		
FL (mm)	n	*Median bias (95% CI) (days)*	*± 14 days (%)*	*Preterm (%)*	*Post-term (%)*
11–22	1981	0.78 (0.07 to 0.90)	84.8	4.2	5.5
23–26	7402	0.49 (0.29 to 0.49)	86.6	3.5	5.4
27–29	14 524	2.28 (1.98 to 2.28)	86.2	3.1	7.2
30–42	16 082	3.98 (3.98 to 4.39)	83.0	2.5	11.8
11–42	39 989	2.61 (2.49 to 2.72)	84.9	3.0	8.6

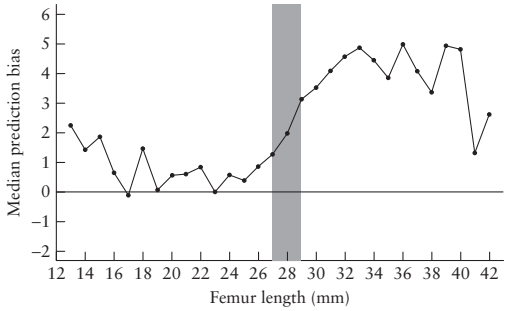


Figure 4 Median bias for the Bergen–2004 model related to different measurements of femur length (FL). ■, Central FL area (week 18).

the predicted date of delivery varied by up to 5 days, depending solely on the fetal size at the time of the examination.

In the analysis of the difference between applying the mean of three or the longest of three FL measurements, we found the overall disagreement to be 0.5 mm; the effect of this amounts to approximately minus 0.8 days (data not shown).

DISCUSSION

In a population of approximately 40 000 pregnancies, we evaluated two sample-based Norwegian models for term prediction and compared the results with predictions from a corresponding German model. We found that the models had a substantial variation in the quality of their predictions resulting in a bias, the extent of which depended on the fetal size at the time of the ultrasound examination.

The median is a robust parameter for evaluating the left-skewed birth residual distribution[21,25]. Our primary evaluation measure was the median bias of each prediction model. The bias measures how well a model is calibrated to the evaluation population – the systematic error in prediction. Therefore, the bias is appropriate when comparing the different models in our population. Precision is the spread of the residual distribution – the random variation around the median prediction value, regardless of bias. Precision is relevant when comparing the predictive quality of different fetal measurements, such as BPD versus FL, but not when comparing different models based on the same measurement[23]. Thus, we did not evaluate precision here, except through the secondary measures, which are influenced by both bias and precision.

There are different reasons for the biases identified in this study. The strict selection of relatively few pregnant women with 'certain' LMP data, an assumed pregnancy length of 282 days added to the 'artificial' LMP estimated from the fetal size, and the fact that only a few sonographers took all the measurements, are factors known to produce systematic biases in the traditional models' predictions[21,25]. In fact, these methods primarily estimate a date for the LMP, with the EDD as a secondary outcome.

The relatively small number of women included has other effects. To obtain smooth prediction curves, a (fractional) polynomial regression model is frequently applied[14]. The downside of this method is the inherent lack of flexibility, which may cause biases, particularly at the endpoints. Additionally, a source of substantial bias is the uniform selection of pregnancies over the inclusion range, intended to ensure an even coverage of all gestational weeks, resulting in a sample distribution that differs greatly from the population of routine examinations[21].

The trend of a fetal size-dependent bias in the Trondheim–1984 model was also observed by Tunón *et al.*[9] in 1998, who concluded that the model needed adjustment. However, as also shown in the present study, for predictions based on BPD measurements of 45–54 mm, corresponding to 18.5–21 weeks' gestation, the bias was negligible. This illustrates the problem of selection bias when the method is applied outside the normal time span for routine examinations or to women other than those who have been highly selected. The Bergen–2004 model was developed 20 years later, but was based on the same principles as Trondheim–1984, evidently generating many of the same problems[14]. Bergen–2004 also performs satisfactorily in a restricted scanning period: in weeks 16–18 for the BPD-based, and in weeks 15–18 for the FL-based predictions. As shown in Figure 4, the median bias of the FL-based predictions after week 20 extends to as much as 5 days, expanding the proportion of post-term pregnancies to nearly 12%. Hansmann–1976 had the greatest span of median bias (6 days), with both too early and too late predictions, but

in the main BPD-prediction region the biases were quite similar to those of Bergen–2004.

A system with high-quality term predictions is a prerequisite for epidemiological surveillance and evaluating various aspects of perinatal outcome. A bias of the predictions in either direction has consequences, especially regarding the management of post-term pregnancies or threatening extremely preterm deliveries. In particular, the scheduling of induction of labor in post-term pregnancies should be based on a precise EDD to avoid their under- or overestimation.

Irrespective of the applied prediction model, recognizing early intrauterine growth restriction is a challenge when dating pregnancies with ultrasonography, because the biological variation in growth is being 'reset' to a mean level for all fetuses. Significant discrepancies between reliable LMP-based and ultrasound-based predicted dates need to be acknowledged.

In this study we used data collected over a period of 19 years. For the newer prediction model developed in 2004 we therefore included a correction for time period[21], as the ultrasound measurements taken earlier may have been influenced by beam width differences in older ultrasound machines[22,26]. This correction for time improves all the Bergen–2004 results, albeit marginally. The axially measured BPD is less influenced by differences in measurement technique and quality of the ultrasound equipment than is the nearly horizontally measured FL.

During the years of data collection, Trondheim–1984 was the only dating method in use at the hospital and was thus the basis for clinical pregnancy management. It is reasonable to ask whether this might skew the results in favor of one model. The post-term inductions are the only relevant events to be considered, but because all these inductions took place from 12 or more days past the EDD predicted by both models, the median biases were not affected. Bergen–2004 predicts term earlier than does Trondheim–1984 for all BPD values; moreover, elective Cesarean sections and all other inductions of labor were excluded. It is therefore unlikely that pregnancy management markedly influenced even the post-term percentages.

The results and conclusions are closely related to the composition of the study population, which in this case reflects the Norwegian practice of routine ultrasound examination at around 17–19 weeks' gestation. Altering the population distribution would, for purely statistical reasons, produce different biases. Consequently, this study does not provide information about the quality of first- vs. second-trimester examinations.

In summary, this study has shown inherent weaknesses in old models for term prediction, as they proved to have varying biases that reduced the quality of the EDD-predictions and over- or underestimated the rates of preterm and post-term pregnancies. The selection criteria of the sample-based models are important in producing these biases. Therefore, the more robust population-based approach to all categories of fetal biometry charts[21,25,27,28] should probably supersede the traditional models.

ACKNOWLEDGMENT

We would like to thank Nancy Lea Eik-Nes for revising the manuscript.

REFERENCES

1. Markestad T, Kaaresen PI, Rønnestad A, Reigstad H, Lossius K, Medbo S, Zanussi G, Engelund IE, Skjaerven R, Irgens LM. Early death, morbidity, and need of treatment among extremely premature infants. *Pediatrics* 2005; **115**: 1289–1298.
2. Heimstad R, Romundstad PR, Eik-Nes SH, Salvesen KA. Outcomes of pregnancy beyond 37 weeks of gestation. *Obstet Gynecol* 2006; **108**: 500–508.
3. Campbell S. The prediction of fetal maturity by ultrasonic measurement of the biparietal diameter. *J Obstet Gynaecol Br Commonw* 1969; **76**: 603–609.
4. Gardosi J, Geirsson RT. Routine ultrasound is the method of choice for dating pregnancy. *Br J Obstet Gynaecol* 1998; **105**: 933–936.
5. Persson PH. Ultrasound dating of pregnancy – still controversial? *Ultrasound Obstet Gynecol* 1999; **14**: 9–11.
6. Tunón K, Eik-Nes SH, Grøttum P. A comparison between ultrasound and a reliable last menstrual period as predictors of the day of delivery in 15 000 examinations. *Ultrasound Obstet Gynecol* 1996; **8**: 178–185.
7. Saltvedt S, Almstrom H, Kublickas M, Reilly M, Valentin L, Grunewald C. Ultrasound dating at 12–14 or 15–20 weeks of gestation? A prospective cross-validation of established dating formulae in a population of *in-vitro* fertilized pregnancies randomized to early or late dating scan. *Ultrasound Obstet Gynecol* 2004; **24**: 42–50.
8. Taipale P, Hiilesmaa V. Predicting delivery date by ultrasound and last menstrual period in early gestation. *Obstet Gynecol* 2001; **97**: 189–194.
9. Tunón K, Eik-Nes SH, Grøttum P. The impact of fetal, maternal and external factors on prediction of the day of delivery by the use of ultrasound. *Ultrasound Obstet Gynecol* 1998; **11**: 99–103.
10. Hutchon DJ, Ahmed F. Naegele's rule: a reappraisal. *BJOG* 2001; **108**: 775.
11. Eik-Nes SH, Grøttum P. Graviditetskalenderen Snurra. ScanMed A/S, Drammen, Norway, 1983.
12. Johnsen SL, Rasmussen S, Sollien R, Kiserud T. Fetal age assessment based on ultrasound head biometry and the effect of maternal and fetal factors. *Acta Obstet Gynecol Scand* 2004; **83**: 716–723.
13. Johnsen SL, Rasmussen S, Sollien R, Kiserud T. Fetal age assessment based on femur length at 10–25 weeks of gestation, and reference ranges for femur length to head circumference ratios. *Acta Obstet Gynecol Scand* 2005; **84**: 725–733.
14. Altman DG, Chitty LS. New charts for ultrasound dating of pregnancy. *Ultrasound Obstet Gynecol* 1997; **10**: 174–191.
15. Hansmann M. Ultraschallbiometrie im II. und III. Trimester des Schwangerschaft. *Gynäkologe* 1976; **9**: 133–155.
16. Campbell S, Thoms A. Ultrasound measurement of the fetal head to abdomen circumference ratio in the assessment of growth retardation. *Br J Obstet Gynaecol* 1977; **84**: 165–174.
17. Goldstein RB, Filly RA, Simpson G. Pitfalls in femur length measurements. *J Ultrasound Med* 1987; **6**: 203–207.
18. Hadlock FP, Harrist RB, Deter RL, Park SK. Fetal femur length as a predictor of menstrual age: sonographically measured. *AJR Am J Roentgenol* 1982; **138**: 875–878.
19. Kurmanavicius J, Wright EM, Royston P, Zimmermann R, Huch R, Huch A, Wisser J. Fetal ultrasound biometry: 2.

Abdomen and femur length reference values. *Br J Obstet Gynaecol* 1999; **106**: 136–143.
20. Rosati P, Guariglia L, Capelli G. A new mathematical formula for predicting long bone length in early pregnancy. *Ultrasound Obstet Gynecol* 2002; **19**: 184–189.
21. 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.
22. Økland I, Bjåstad TG, Johansen TF, Gjessing HK, Grøttum P, Eik-Nes SH. OC14.06: Femur length measurements are influenced by beam width. *Ultrasound Obstet Gynecol* 2009; **34** (Suppl. 1): 27.
23. Gjessing HK, Grøttum P. Accuracy of second trimester fetal head circumference and biparietal diameter for predicting the time of spontaneous birth. *J Perinat Med* 2007; **35**: 350–351; author reply 351–352.
24. R Development Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna, 2008; http://www.R-project.org. [Accessed 17 December 2008].
25. Salomon LJ, Pizzi C, Gasparrini A, Bernard JP, Ville Y. Prediction of the date of delivery based on first trimester ultrasound measurements: an independent method from estimated date of conception. *J Matern Fetal Neonatal Med* 2010; **23**: 1–9.
26. Jago JR, Whittingham TA, Heslop R. The influence of ultrasound scanner beam width on femur length measurements. *Ultrasound Med Biol* 1994; **20**: 699–703.
27. Sahota DS, Kagan KO, Lau TK, Leung TY, Nicolaides KH. Customized birth weight: coefficients and validation of models in a UK population. *Ultrasound Obstet Gynecol* 2008; **32**: 884–889.
28. Verburg BO, Steegers EA, De Ridder M, Snijders RJ, Smith E, Hofman A, Moll HA, Jaddoe VW, Witteman JC. New charts for ultrasound dating of pregnancy and assessment of fetal growth: longitudinal data from a population-based cohort study. *Ultrasound Obstet Gynecol* 2008; **31**: 388–396.

Paper II

Ultrasound Obstet Gynecol 2011; **37**: 207–213
Published online 12 January 2011 in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/uog.7728

A new population-based term prediction model vs. two traditional sample-based models: validation on 9046 ultrasound examinations

I. ØKLAND*†, H. K. GJESSING‡§, P. GRØTTUM¶, T. M. EGGEBØ† and S. H. EIK-NES*#

**National Center for Fetal Medicine, Women and Children's Center, St Olavs University Hospital, Trondheim, Norway; †Department of Obstetrics and Gynecology, Stavanger University Hospital, Stavanger, Norway; ‡Division of Epidemiology, Norwegian Institute of Public Health, Oslo, Norway; §Department of Public Health and Primary Health Care, University of Bergen, Bergen, Norway; ¶Section of Medical Informatics, University of Oslo, Oslo, Norway; #Norwegian University of Science and Technology, Trondheim, Norway*

KEYWORDS: date of delivery; fetal measurements; gestational age; post-term pregnancy; routine ultrasound examination; second trimester

ABSTRACT

Objectives *To compare results of predictions of date of delivery from a new population-based model with those from two traditional regression models.*

Methods *We included 9046 fetal biparietal diameter (BPD) measurements and 8776 femur length (FL) measurements from the routine ultrasound examinations at Stavanger University Hospital between 2001 and 2007. The prediction models to be validated were applied to the data, and the resulting predictions were compared with the actual time of the subsequent deliveries. The primary measure was the median bias (the difference between the true and the predicted date of delivery), calculated for each method, for the study population as a whole and for three subgroups of BPD/FL measurements. We also assessed the proportion of births within ± 14 days of the predicted day, and rates of preterm and post-term deliveries, which were regarded as secondary measures.*

Results *For the population-based model, the median bias was −0.15 days (95% confidence interval (CI), −0.43 to 0.12) for the BPD-based, and −0.48 days (95% CI, −0.86 to −0.46) for the FL-based predictions, and both biases were stable over the inclusion ranges. The biases of the traditional regression models varied, depending on the fetal size at the time of the examination; the extremes were −3.2 and + 4.5 days for the BPD-based, and −1.0 and + 5.0 days for the FL-based predictions.*

Conclusions *The overall biases, as well as the biases for the subgroups, were all smaller with the population-based model than with the traditional regression models, which exhibited substantial biases in some BPD and FL subcategories. For the population-based model, the FL-based predictions were in accordance with the BPD-based predictions. Copyright © 2011 ISUOG. Published by John Wiley & Sons, Ltd.*

INTRODUCTION

Most pregnancies in western countries are dated by a first- or second-trimester ultrasound examination. Since Campbell introduced the first models for estimating fetal age from biparietal diameter (BPD) measurements[1], discussion on how to predict the 'correct' date of confinement has been incessant[2,3]. During the 1990s several studies showed that dating by the last menstrual period (LMP) was unreliable for a considerable proportion of pregnant women[4,5]. An agreement emerged that ultrasound dating was superior to that based on Naegele's rule[6–8]. In recent years, the debate has been when, what and how to measure[8–10], and which prediction method to use[11–13]. BPD-based and femur length (FL)-based estimates remain standard in second-trimester pregnancy dating[10].

Our clinical protocols concerning the management of preterm and post-term deliveries are based on risk estimates for fetuses at different gestational ages[14,15]. Evaluations of protocols and comparisons of outcome are inevitably linked to a uniform and correct pregnancy dating, and the ongoing discussions about when to induce labor in post-term pregnancies[16,17] are meaningful only when based on standardized ultrasound-based dating methods[2,18].

Correspondence to: Dr I. Økland, Department of Obstetrics and Gynecology, Stavanger University Hospital, P.O. Box 8100, N-4068 Stavanger, Norway (e-mail: iok@lyse.net)

Accepted: 26 May 2010

Until recently, the methodological basis for obstetric wheels for term prediction has been regression of the LMP-based gestational age on fetal ultrasound measurements. The models thus calculate an 'artificial' LMP[19], and 280–282 days has to be added to obtain the estimated date of delivery (EDD). These methods primarily estimate an LMP and the subsequent conception, with the EDD as a secondary objective. To construct such models, a highly selected sample of pregnant women with 'certain' dates of LMP must be recruited[20]. Consequently, the selected women are not representative of the 'typical' expectant mother at a routine examination, and the prediction model is still influenced by the LMP[19]. Therefore, because the traditional approach runs a risk of producing biased predictions[13,21,22], a population-based term prediction model was developed[23], predicting the day of delivery directly. This model circumvents the LMP and uses data from a population-based clinical registry.

The purpose of this study was to determine the validity of three different methods for term prediction – two traditional and the new population-based model – on data from a clinical database.

SUBJECTS AND METHODS

Subjects

The women included in this study had their routine fetal ultrasound examination and subsequent delivery at Stavanger University Hospital between January 2001 and November 2006. In accordance with the Norwegian practice of routine ultrasound examination, the scan was scheduled to take place at between 17 and 20 completed weeks according to the LMP and/or an early clinical examination. The majority of the examinations were carried out in weeks 17–19. We included all women with singleton pregnancies whose fetus had a BPD in the range of 38–60 mm at the routine scan, and who gave birth to a liveborn child without anomalies after 23 completed weeks.

A total of 10 193 women were initially considered for inclusion, of whom 1147 were excluded: 704 were excluded because of induction of labor for reasons other than post-term pregnancy, 442 had an elective Cesarean section prior to the start of labor and one had missing information, leaving 9046 women in the study. Women induced for post-term pregnancy ($n = 301$) were not excluded.

Prediction models

Three different models for prediction of the date of delivery were validated in this study. From 1984 until recently, the obstetric wheel 'Snurra' (referred to here as 'Trondheim–1984')[24] has been the only dating method in use in Norway, predicting the date of delivery from second-trimester BPD measurements of between 38 and 60 mm only. The model was constructed from a population of 90 women with anticipated normal pregnancies, all of whom were carefully selected regarding menstrual history. The women were included in a prospective, longitudinal study, and measurements were taken from each fetus approximately 10 times. Fourth-order polynomial regression analysis was used to establish the curves.

A newly developed obstetric wheel 'Terminhjulet' ('Bergen–2004')[25,26] is based on the same principles[20] as Trondheim–1984, but uses a newer data sample and fractional polynomial regression analysis. This model was constructed from a prospective, cross-sectional study of 650 healthy women with regular menstrual periods and singleton, uncomplicated pregnancies. With the Bergen–2004 model the date of delivery may be predicted from BPD (14–60 mm), FL (2–44 mm) or from head circumference (50–134 mm) measurements.

Both of these models are traditional and sample based, and based on limited reference material. The prediction model 'eSnurra' ('Trondheim–2007')[23] is an implementation of the new population-based model for direct prediction of the date of delivery, based on second-trimester fetal measurements from an unselected population of 37 000 singleton pregnancies. From these measurements, the median remaining time of pregnancy was computed, using a local linear quantile regression model. Trondheim–2007 predicts date of delivery from BPD (25–60 mm) or from FL (11–42 mm) measurements.

The BPD prediction tables for Trondheim–1984 are available from this prediction model[24]. For Bergen–2004, the BPD- and FL-tables in the articles show the computed gestational age, but the accompanying published formulae[25,26] are in error. The intended formulae[22] have been provided by S. L. Johnsen (pers. comm.).

For both the Trondheim–1984 and the Bergen–2004 models, EDD is found by adding 282 days to the inferred LMP-date. For Trondheim–2007, remaining time is calculated from the published tables[23]. The model advocates a correction for time periods, and the correction values for the latest time period (1999–2004) have been included.

Ultrasound examinations

The routine fetal ultrasound examinations were performed by seven formally trained midwives, and the data were prospectively registered in a computer database.

All clinical problems were primarily managed according to the Trondheim–1984 prediction model; post-term pregnant women were scheduled for induction of labor after 13 days past EDD (corresponding to ≥ 295 days), and induction for clinical problems possibly related to post-term pregnancy was not acknowledged as such until 10 days past EDD (≥ 292 days).

BPD was measured as described by Campbell and Thoms[27], through a horizontal section of the fetal head at the level of the cavum septi pellucidi, from the outer to the outer contours of the parietal bones. The mean of three BPD measurements was used for the calculation. FL was measured in accordance with the method described

by Goldstein *et al.*[28], with the femoral diaphysis in a longitudinal section. The ossified part of the femoral diaphysis was measured three times and the longest measurement was used[29–31]. Both BPD and FL were rounded to the nearest millimeter.

The Bergen–2004 model used the mean of three FL measurements, rather than adhering to the international practice of using the longest of three FL measurements[29–31]; this consideration has been addressed in a previous study[22].

The following ultrasound scanners were used: Brüel & Kjær Medical System 3535 and 2102 (Copenhagen, Denmark), Hitachi EUB–5500, EUB–6000 and EUB–6500 (Kashiwa, Japan) and Voluson 730 Expert (Zipf, Austria) with 3.5–7.5-MHz multifrequency curvilinear transducers.

Statistical methods

The three prediction models were applied to data from the ultrasound examinations and from the subsequent deliveries. The 9046 measurements with BPD in the range 38–60 mm were used to predict the date of delivery according to all three models. The 8776 (of the 9046) measurements where an FL measurement in the range 23–42 mm was available were used with Bergen–2004 and Trondheim–2007. Our primary measure for assessing prediction quality was the median bias, defined as the difference, in days, between the true and the predicted date of delivery. A negative bias means that the birth occurred before the predicted date, resulting in an apparent decrease in the number of births defined as post-term. A positive bias means that birth took place later than the day predicted by the model. A systematic positive bias will lead to a seemingly inflated post-term rate. To measure this effect of misclassification, we also computed secondary measures, i.e., the proportion of births within ± 14 days of the EDD, and proportions of preterm (24 days or more before EDD) and post-term (14 days or more after EDD) deliveries. An altered median bias will affect the rate of deliveries defined as post-term and the proportion of births within ± 14 days more than it will the rate of deliveries defined as preterm.

Bias may vary over the range of inclusion, and a variable bias, depending on the fetal size at the time of the examination, may be overlooked if only the overall bias is calculated[32]. Therefore, for each method, median bias and secondary measures were calculated for the study population as a whole and for three subgroups of BPD/FL measurements. The three models were used with the same study population, therefore the secondary measures merely indicate the effects of the size and shifting of the bias.

P-values for testing a non-zero median bias were computed using permutation tests with 2000 permutations[33]. Confidence intervals for the secondary measures were computed using the Wilson method in the 'binom' package in R[34]. All analyses and graphics were produced in the *R* statistical programming environment[35].

RESULTS

The median age of the women in the study population was 29.3 years; 41.1% were primiparous and labor was induced in 9.9%, not including inductions as a result of prelabor rupture of membranes. There were approximately 4000 births per year at the hospital, and the Cesarean section rate was around 11% during the study period.

BPD-based prediction

Table 1 shows the median biases and the values of the secondary measures (the proportion of births within ± 14 days and rates of preterm and post-term deliveries), when date of delivery is predicted with each of the three models, for three different groups with BPD ranges corresponding to a gestational age below 18 weeks (38–43 mm), week 18 (44–46 mm) and above 18 weeks (47–60 mm), and for the study group as a whole. The secondary measures demonstrate how variations in the median bias among subgroups and prediction models affect particularly the rate of apparently post-term pregnancies, due to the left-skewed birth distribution curves with the steep down-slope on the right side (Figure 1). A median bias of −2.75 days for Trondheim–1984 in one BPD-subgroup gives a post-term rate of 2%, while a similar, but inverse, bias of 2.94 days for Bergen–2004 in another subgroup results in 10.7% of the pregnancies being classified as post-term.

Figure 1 shows the birth distribution histograms relative to the date of delivery predicted by the three different models for the study group as a whole, and it illustrates the systematic shift of the distribution of birth residuals resulting from the varying size of the median biases. The median biases for the traditional models (Trondheim–1984 and Bergen–2004) varied between −3.2 and +4.5 days over the BPD inclusion range, while it was stable for the new population-based model (Trondheim–2007), essentially within ± 1 day, as shown in Figure 2.

FL-based prediction

The two models Bergen–2004 and Trondheim–2007 were applied to the FL measurements. The overall distribution curves of birth relative to the date of delivery predicted by the two different models are shown in Figure 3, which illustrates the systematic shift of the distribution of birth residuals resulting from the varying size of the median bias.

In Table 2 the median biases, the proportion of births within ± 14 days and rates of preterm and post-term deliveries are presented for three different groups with FL ranges corresponding to a gestational age of less than 18 weeks (23–26 mm), week 18 (27–29 mm) and greater than 18 weeks (30–42 mm), and for the study group as a whole.

Table 1 Term prediction from biparietal diameter (BPD) measurement with the three different models: median bias, proportion of births within ± 14 days and proportion of preterm and post-term deliveries for the three BPD-range groups and for the study group as a whole

			Primary measure		*Secondary measures*		
BPD (mm)	n	*Method*	*Median bias (days (95% CI))*	P	*Birth within ± 14 days (% (95% CI))*	*Preterm delivery (% (95% CI))*	*Post-term delivery (% (95% CI))*
38–43	1841	Trondheim–1984	−2.75 (−3.29 to −2.26)	< 0.01	85.0 (83.2–86.5)	5.6 (4.7–6.8)	2.0 (1.4–2.7)
		Bergen–2004	0.54 (0.43 to 1.43)	< 0.01	85.7 (84.0–87.2)	4.1 (3.3–5.1)	5.9 (4.9–7.0)
		Trondheim–2007	0.56 (−0.15 to 0.85)	0.02	85.5 (83.8–87.0)	4.2 (3.4–5.2)	5.6 (4.6–6.7)
44–46	3238	Trondheim–1984	−1.81 (−2.17 to −1.44)	< 0.01	86.2 (84.9–87.3)	5.5 (4.8–6.4)	2.6 (2.1–3.2)
		Bergen–2004	0.88 (0.74 to 1.74)	< 0.01	85.2 (83.9–86.4)	4.4 (3.7–5.1)	5.8 (5.1–6.7)
		Trondheim–2007	−0.43 (−0.74 to 0.26)	0.01	86.1 (84.9–87.3)	4.8 (4.1–5.6)	4.0 (3.3–4.7)
47–60	3967	Trondheim–1984	−0.77 (−1.11 to −0.41)	< 0.01	86.5 (85.4–87.5)	4.9 (4.3–5.6)	4.4 (3.8–5.1)
		Bergen–2004	2.94 (2.23 to 3.04)	< 0.01	83.3 (82.1–84.4)	3.2 (2.7–3.8)	10.7 (9.8–11.7)
		Trondheim–2007	−0.30 (−0.70 to 0.12)	0.11	86.2 (85.1–87.3)	4.2 (3.7–4.9)	4.8 (4.2–5.5)
38–60	9046	Trondheim–1984	−1.57 (−1.80 to −1.35)	< 0.01	86.1 (85.3–86.8)	5.3 (4.8–5.8)	3.2 (2.9–3.6)
		Bergen–2004	1.74 (1.54 to 2.04)	< 0.01	84.4 (83.7–85.2)	3.8 (3.4–4.2)	8.0 (7.4–8.6)
		Trondheim–2007	−0.15 (−0.43 to 0.12)	0.29	86.0 (85.3–86.7)	4.4 (4.0–4.9)	4.7 (4.2–5.1)

Table 2 Term prediction from femur length (FL) measurement with two different models: median bias, proportion of births within ± 14 days and proportion of preterm and post-term deliveries for three FL-range groups and for the study group as a whole

			Primary measure		*Secondary measures*		
FL (mm)	n	*Method*	*Median bias (days (95% CI))*	P	*Birth within ± 14 days (% (95% CI))*	*Preterm delivery (% (95% CI))*	*Post-term delivery (% (95% CI))*
23–26	1802	Bergen–2004	0.60 (−0.16 to 0.84)	< 0.01	84.8 (83.1–86.4)	5.1 (4.2–6.2)	5.0 (4.1–6.1)
		Trondheim–2007	−0.27 (−0.54 to 0.47)	0.14	84.9 (83.1–86.4)	5.5 (4.6–6.7)	3.9 (3.1–4.9)
27–29	3463	Bergen–2004	1.26 (0.91 to 1.26)	< 0.01	86.8 (85.7–87.9)	4.1 (3.5–4.8)	5.1 (4.4–5.9)
		Trondheim–2007	−0.86 (−1.26 to −0.48)	< 0.01	87.5 (86.3–88.5)	4.7 (4.1–5.5)	2.4 (2.0–3.0)
30–42	3511	Bergen–2004	3.52 (3.06 to 4.02)	< 0.01	81.9 (80.5–83.1)	3.2 (2.7–3.9)	12.2 (11.1–13.3)
		Trondheim–2007	−0.46 (−0.74 to −0.44)	0.04	85.3 (84.1–86.5)	4.4 (3.8–5.1)	4.1 (3.5–4.8)
23–42	8776	Bergen–2004	1.91 (1.60 to 2.06)	< 0.01	84.4 (83.6–85.2)	3.9 (3.6–4.4)	7.9 (7.4–8.5)
		Trondheim–2007	−0.48 (−0.86 to −0.46)	< 0.01	86.1 (85.3–86.8)	4.8 (4.3–5.2)	3.4 (3.0–3.8)

The median bias for the traditional model, Bergen–2004, increased with increasing FL, while for the population-based model, Trondheim–2007, it was stable within ± 1 day over the full inclusion range of FL, as shown in Figure 4.

The study population of approximately 9000 pregnancies resulted in significant *P*-values, indicating median biases significantly different from zero, even for clinically insignificant biases of less than 0.5 days (Tables 1 and 2). Both Trondheim–1984 and Bergen–2004 had significant median biases for all subgroups and for the study population as a whole, while the biases of Trondheim–2007 were for several groups not statistically significantly different from zero.

DISCUSSION

In this study, two traditional sample-based models and one new population-based model for the prediction of date of delivery were validated in a population of 9046 pregnancies. We found that the new population-based model made reliable predictions from second-trimester measurements of BPD or FL. In contrast, both traditional models had substantial variations in the quality of their predictions, resulting in a bias. The extent of the bias depended on the fetal size at the time of the ultrasound examination. The direction and size of the biases were equivalent to those found for the same sample-based models in a recent study of the models used with a sample of 41 343 ultrasound examinations[22]. These reproducible biases of the sample-based models show that inherent weaknesses are built into the traditional term prediction methods.

The median is a robust parameter for evaluating the markedly left-skewed residual distribution of births[13,23]. Our primary measure of predictive quality is the median bias of each prediction model, which measures the systematic error in prediction: the calibration of the model to the study population. Therefore, the bias is relevant when comparing three different models in our population[22]. The precision is the random variation around the median prediction value, and is not appropriate for use when comparing different models based on the same fetal measurement[32]. The secondary measures of deviation, shown in Tables 1 and 2, are influenced by both bias and precision.

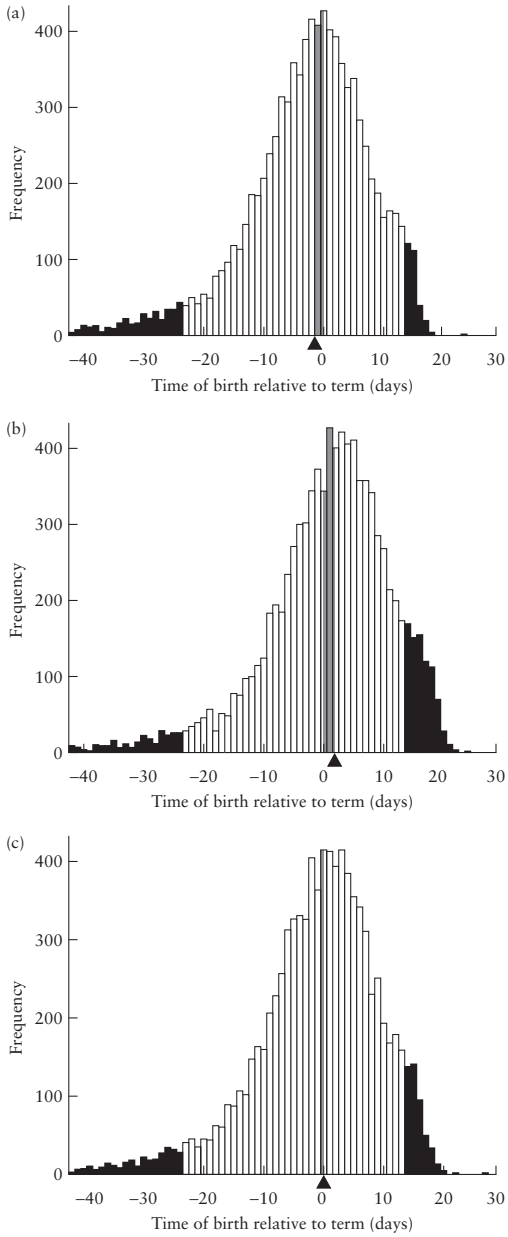


Figure 1 Birth distribution histograms showing the median time of birth (▲) relative to the predicted term (0) from 9046 measurements of biparietal diameter for (a) the Trondheim–1984 model, (b) the Bergen–2004 model and (c) the Trondheim–2007 model. ▒, Days of median bias; ■, preterm and post-term deliveries.

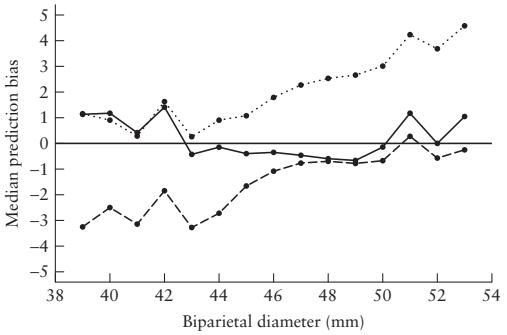


Figure 2 Median biases for the three models (Trondheim–1984 (- - - -), Bergen–2004 (·····) and Trondheim–2007 (——)) related to different biparietal diameter measurements.

There are a number of possible reasons for the biases observed in the traditional models. The most 'trivial' is the assumed total length of pregnancy in the population, in our case 282 days. Since this number is used in every single prediction, a systematic bias will appear if this number does not correspond to the actual median length[13]. Another problem is the careful selection of pregnant women that leads to a prediction model based on a 'hyper-normal' population; when applied to a population of women with unreliable or missing LMP data, biases may appear. The traditional models are also often based on measurements performed by only one or two dedicated sonographers. Many authors regard this as an important cause of the systematic bias seen in traditional prediction models[13,23,30].

The notably smaller biases following the use of the population-based model have several explanations. By estimating the remaining time of pregnancy through a direct regression of time to delivery on fetal size, this model is independent of the LMP. Interestingly, since the model is calibrated to the correct median remaining time, knowledge of the total length of pregnancy is unnecessary for term prediction, but is still essential for the estimation of gestational age[36] – the exact opposite of the LMP-based models. This model was based on a large population and measurements were made by many sonographers. This makes the population-model robust against selection bias[13,19,23]. Finally, the large sample sizes make it feasible to use, for instance, nonparametric quantile regression, which is more flexible and robust than is polynomial regression[36].

The quality of the predictions for the traditional models in our study was not uniform. Over the range of inclusion, the median biases varied widely between the two models; they generally predicted delivery date too early or too late. Additionally, the bias for both models varied with the time of the examination, producing a slope in the median bias curve for both the BPD- and the FL-based predictions, as seen in Figures 2 and 4. For the same value of BPD, the difference in the EDD between the two models amounted to as much as 4 days. The secondary measures of our evaluation were, of course, directly influenced by the bias of each model.

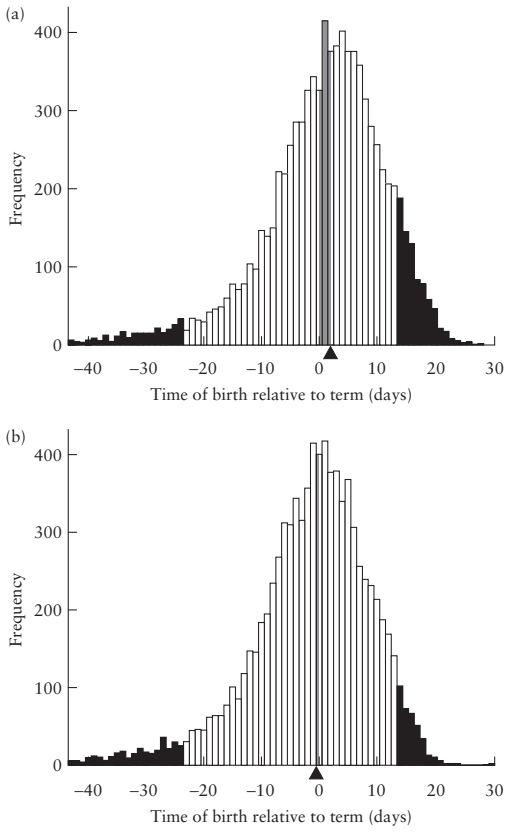


Figure 3 Birth distribution histograms showing the median time of birth (▲) relative to the predicted term (0) from 8776 measurements of femur length for (a) the Bergen–2004 model and (b) the Trondheim–2007 model. ▒, Days of median bias; ■, preterm and post-term deliveries.

Post-term induction routines should be based on reliable estimates of date of delivery; a few days of displacement in either direction will under- or overestimate the number of true post-term pregnancies, resulting in increased risk for post-mature fetuses not being recognized as such, or iatrogenic, 'unnecessary' inductions of labor in women who have barely passed their EDD. Post-term pregnancies are associated with a small, but not negligible increase in the risk of mortality, probably after 41 weeks[37]. Therefore, a precise term prediction is an important factor for true age-related risk assessment[18].

The incidence of post-term deliveries (≥ 294 days of gestation) varies in different studies, but is generally below 4% when EDD is calculated from ultrasound parameters[9,16,38]. The gradually adopted policy of earlier post-term induction[16] implies a two- to three-fold increase in induction rates, depending on the chosen cut-off. The clinical problems associated with monitoring and/or inducing labor in false positive post-term pregnant women

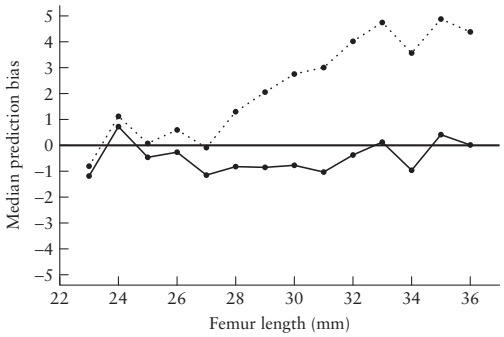


Figure 4 Median biases for the two models (Bergen–2004 (.....) and Trondheim–2007 (——)) related to different femur length measurements.

due to inadequate predictions, put an extra strain on delivery units[16], independent of the discussions on when to induce in post-term pregnancy.

During the years of data collection, Trondheim–1984 was the only dating method in use at the hospital and was thus the basis for all clinical pregnancy management. Questions as to whether this might favor the results of this model or increase the biases of the other two models in the study may be raised. Such a bias could be introduced by the fact that a pregnancy might be considered overdue according to the standard dating method, but not by the others. Trondheim–1984 predicts term later than do the other models for all BPD measures. Therefore, the post-term pregnancies in our study would have been even more post-term according to Bergen–2004 and Trondheim–2007. The included post-term inductions took place well beyond the date predicted by all three models. All other inductions of labor and all elective Cesarean sections performed prior to the start of labor were excluded, because the scheduling of these events is influenced by the EDD.

This study has shown that the new population-based model had the most accurate term predictions, being superior to the two traditional methods. The overall biases, as well as the biases for the different subgroups, were all smaller with this model, and the BPD- and FL-based predictions were nearly equivalent. Population-based term prediction may be the dating method of choice for years to come.

ACKNOWLEDGMENTS

Leif K. Gjessing supplied us with the data files. Nancy Lea Eik-Nes revised the manuscript.

REFERENCES

1. Campbell S. The prediction of fetal maturity by ultrasonic measurement of the biparietal diameter. *J Obstet Gynaecol Br Commonw* 1969; **76**: 603–609.

2. Mongelli M, Wilcox M, Gardosi J. Estimating the date of confinement: ultrasonographic biometry versus certain menstrual dates. *Am J Obstet Gynecol* 1996; **174**: 278–281.
3. Rossavik IK, Fishburne JI. Conceptional age, menstrual age, and ultrasound age: a second-trimester comparison of pregnancies of known conception date with pregnancies dated from the last menstrual period. *Obstet Gynecol* 1989; **73**: 243–249.
4. Tunón K, Eik-Nes SH, Grøttum P. A comparison between ultrasound and a reliable last menstrual period as predictors of the day of delivery in 15 000 examinations. *Ultrasound Obstet Gynecol* 1996; **8**: 178–185.
5. Nguyen TH, Larsen T, Engholm G, Møller H. Evaluation of ultrasound-estimated date of delivery in 17 450 spontaneous singleton births: do we need to modify Naegele's rule? *Ultrasound Obstet Gynecol* 1999; **14**: 23–28.
6. Gardosi J, Geirsson RT. Routine ultrasound is the method of choice for dating pregnancy. *Br J Obstet Gynaecol* 1998; **105**: 933–936.
7. Neilson JP. Ultrasound for fetal assessment in early pregnancy. *Cochrane Database Syst Rev* 1998 Issue 4; DOI: 10.1002/14651858.
8. Persson PH. Ultrasound dating of pregnancy–still controversial? *Ultrasound Obstet Gynecol* 1999; **14**: 9–11.
9. Taipale P, Hiilesmaa V. Predicting delivery date by ultrasound and last menstrual period in early gestation. *Obstet Gynecol* 2001; **97**: 189–194.
10. Verburg BO, Steegers EA, De Ridder M, Snijders RJ, Smith E, Hofman A, Moll HA, Jaddoe VW, Witteman JC. New charts for ultrasound dating of pregnancy and assessment of fetal growth: longitudinal data from a population-based cohort study. *Ultrasound Obstet Gynecol* 2008; **31**: 388–396.
11. Saltvedt S, Almstrom H, Kublickas M, Reilly M, Valentin L, Grunewald C. Ultrasound dating at 12–14 or 15–20 weeks of gestation? A prospective cross-validation of established dating formulae in a population of *in-vitro* fertilized pregnancies randomized to early or late dating scan. *Ultrasound Obstet Gynecol* 2004; **24**: 42–50.
12. Mongelli M, Yuxin NG, Biswas A, Chew S. Accuracy of ultrasound dating formulae in the late second-trimester in pregnancies conceived with *in-vitro* fertilization. *Acta Radiol* 2003; **44**: 452–455.
13. Salomon LJ, Pizzi C, Gasparrini A, Bernard JP, Ville Y. Prediction of the date of delivery based on first trimester ultrasound measurements: an independent method from estimated date of conception. *J Matern Fetal Neonatal Med* 2010; **23**: 1–9.
14. Markestad T, Kaaresen PI, Ronnestad A, Reigstad H, Lossius K, Medbo S, Zanussi G, Engelund IE, Skjaerven R, Irgens LM. Early death, morbidity, and need of treatment among extremely premature infants. *Pediatrics* 2005; **115**: 1289–1298.
15. Heimstad R, Romundstad PR, Eik-Nes SH, Salvesen KA. Outcomes of pregnancy beyond 37 weeks of gestation. *Obstet Gynecol* 2006; **108**: 500–508.
16. Wennerholm UB, Hagberg H, Brorsson B, Bergh C. Induction of labor versus expectant management for post-date pregnancy: is there sufficient evidence for a change in clinical practice? *Acta Obstet Gynecol Scand* 2009; **88**: 6–17.
17. Cleary-Goldman J, Bettes B, Robinson JN, Norwitz E, D'Alton ME, Schulkin J. Postterm pregnancy: practice patterns of contemporary obstetricians and gynecologists. *Am J Perinatol* 2006; **23**: 15–20.
18. Norwitz ER, Snegovskikh VV, Caughey AB. Prolonged pregnancy: when should we intervene? *Clin Obstet Gynecol* 2007; **50**: 547–557.
19. Hutchon DJ, Ahmed F. Naegele's rule: a reappraisal. *BJOG* 2001; **108**: 775.
20. Altman DG, Chitty LS. New charts for ultrasound dating of pregnancy. *Ultrasound Obstet Gynecol* 1997; **10**: 174–191.
21. Tunon K, Eik-Nes SH, Grøttum P. The impact of fetal, maternal and external factors on prediction of the day of delivery by the use of ultrasound. *Ultrasound Obstet Gynecol* 1998; **11**: 99–103.
22. Økland I, Gjessing HK, Grøttum P, Eik-Nes SH. Biases of traditional term prediction models: results from different sample-based models evaluated on 41 343 ultrasound examinations. *Ultrasound Obstet Gynecol* 2010; **36**: 728–734.
23. 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.
24. Eik-Nes SH, Grøttum P. *Graviditetskalenderen Snurra*. ScanMed A/S: Drammen, Norway, 1983.
25. Johnsen SL, Rasmussen S, Sollien R, Kiserud T. Fetal age assessment based on ultrasound head biometry and the effect of maternal and fetal factors. *Acta Obstet Gynecol Scand* 2004; **83**: 716–723.
26. Johnsen SL, Rasmussen S, Sollien R, Kiserud T. Fetal age assessment based on femur length at 10–25 weeks of gestation, and reference ranges for femur length to head circumference ratios. *Acta Obstet Gynecol Scand* 2005; **84**: 725–733.
27. Campbell S, Thoms A. Ultrasound measurement of the fetal head to abdomen circumference ratio in the assessment of growth retardation. *Br J Obstet Gynaecol* 1977; **84**: 165–174.
28. Goldstein RB, Filly RA, Simpson G. Pitfalls in femur length measurements. *J Ultrasound Med* 1987; **6**: 203–207.
29. Hadlock FP, Harrist RB, Deter RL, Park SK. Fetal femur length as a predictor of menstrual age: sonographically measured. *AJR Am J Roentgenol* 1982; **138**: 875–878.
30. Kurmanavicius J, Wright EM, Royston P, Zimmermann R, Huch R, Huch A, Wisser J. Fetal ultrasound biometry: 2. Abdomen and femur length reference values. *Br J Obstet Gynaecol* 1999; **106**: 136–143.
31. Rosati P, Guariglia L, Capelli G. A new mathematical formula for predicting long bone length in early pregnancy. *Ultrasound Obstet Gynecol* 2002; **19**: 184–189.
32. Gjessing HK, Grøttum P. Accuracy of second trimester fetal head circumference and biparietal diameter for predicting the time of spontaneous birth. *J Perinat Med* 2007; **35**: 350–351, author reply 351–352.
33. Good P. *Permutation Tests*. Springer-Verlag: New York, 2000.
34. Agresti A, Coull BA. Approximate is better than 'exact' for interval estimation of binomial proportions. *Am Stat* 1998; **52**: 119–126.
35. R Development Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing: Vienna, Austria, 2008; http://www.R-project.org [Accessed 17 December 2008].
36. Gjessing HK, Grøttum P, Økland I, Eik-Nes SH. Comment and reply on: prediction of the date of delivery based on first trimester ultrasound measurements: an independent method from estimated date of conception. *J Matern Fetal Neonatal Med* 2010; **23**: 944–946, author reply 946–947.
37. Hilder L, Costeloe K, Thilaganathan B. Prolonged pregnancy: evaluating gestation-specific risks of fetal and infant mortality. *Br J Obstet Gynaecol* 1998; **105**: 169–173.
38. Lynch CD, Zhang J. The research implications of the selection of a gestational age estimation method. *Paediatr Perinat Epidemiol* 2007; **21** (Suppl. 2): 86–96.

Paper III

Ultrasound Obstet Gynecol 2011
Accepted manuscript online: 25 Aug 2011. DOI: 10.1002/uog.10081

Advantages of the population-based approach to pregnancy dating demonstrated with results from 23 020 ultrasound examinations

I. ØKLAND[1,2], J. NAKLING[3], H. K. GJESSING[4,5], P. GRØTTUM[6], S. H. EIK-NES[1,7]

[1]*Department of Obstetrics and Gynecology, Stavanger University Hospital, Stavanger*
[2]*National Center for Fetal Medicine, Department of Obstetrics and Gynecology, St Olavs University Hospital, Trondheim*
[3]*Department of Obstetrics and Gynecology, Innlandet Hospital Trust, Lillehammer*
[4]*Division of Epidemiology, Norwegian Institute of Public Health, Oslo*
[5]*Department of Public Health and Primary Health Care, University of Bergen, Bergen*
[6]*Section of Medical Informatics, University of Oslo, Oslo*
[7]*Department of Laboratory Medicine, Children's and Women's Health, Norwegian University of Science and Technology, Trondheim, Norway*

KEYWORDS: Date of delivery, fetal measurements, gestational age, post-term pregnancy, second trimester, term prediction.

ABSTRACT

Objectives *To confirm the results from two previous evaluations of term prediction models in a third population, and to explore why biased predictions are unavoidable with sample-based models.*

Methods *In a study population of 23 020 second-trimester ultrasound examinations, data were prospectively collected and registered over the period 1988–2009. Three different models for ultrasonically estimated date of delivery were applied to the measurements of fetal biparietal diameter (BPD) and femur length (FL), and the resulting term estimations were compared with the actual time of the delivery. The difference between the actual and the predicted date of delivery (the median bias) was calculated for each of the models, for three BPD/FL-measurement subgroups and for the study population as a whole.*

Results *For the population-based model, the median bias was +0.4 days for the BPD-based, and –0.4 days for the FL-based predictions, and the biases were stable over the inclusion ranges. The biases of the traditional models varied with the size of the fetus at examination; median biases were –0.87 and +2.2 days, respectively, with extremes –4.2 and +4.8 days for the BPD-based, and +1.72 with range –0.8 to +4.5 days for the FL-based predictions. The disagreement between these two methods was never less than 2 days.*

Conclusion *This study confirms the results from previous studies; median biases were negligible with term predictions from the population-based model, while those from the traditional models varied substantially. The biases, which have clinical implications, seem inevitable with the sample-based models, which, even when tentatively calibrated, will perform unsatisfactorily.*

INTRODUCTION

In modern pregnancy care, it is recommended to date pregnancies by ultrasound in the first or second trimester[1, 2]. Thus, it is imperative that the dating models are reliable. To assess prediction quality and reveal potential systematic biases the models must be evaluated in large populations. In earlier studies, we evaluated two traditional, sample-based models for term prediction and demonstrated significant and nearly identical biases when the models were applied to different populations[3, 4]. We also validated a population-based model that avoided the biases[4], which generally appear to be model-dependent.

The terms 'assessment of gestational age (GA)' and 'estimation of date of delivery' are considered almost synonymous. However, the calculations in fact concern totally different times in a pregnancy. Traditional models primarily estimate a last menstrual period (LMP) from second-trimester fetal measurements, thus, the estimated date of delivery (EDD) is actually an indirect and secondary issue[3, 5, 6]. Conversely, the new population-based models are constructed from observations of the verified date of delivery, predicting remaining time of pregnancy and EDD directly from first- or second-trimester fetal measurements[5, 7].

The first-trimester screening tests[8, 9] and the management of extremely preterm deliveries close to the limits of viability[10, 11] relate to an accurately calculated GA, while the scheduling of invasive procedures and interventions in later pregnancy and determining when to induce in post-term pregnancies[12, 13], depend on knowledge of a reliable EDD. Therefore, it is still relevant to demonstrate prediction biases in the range of 2–5 days.

Because the EDD is model-dependent, recommending post-term induction practices without a uniform system for pregnancy dating is futile[12, 14, 15]. Moreover, different term prediction routines and induction practices make it impossible to compare important perinatal quality indicators. This has major consequences for comparison of data between countries, regions and even hospitals[12].

The significantly biased term predictions that resulted from applying the sample-based models[3, 4] were avoided with a direct, LMP-independent model[4]. In this study, we extend previous findings to a population examined with slightly different routines and different sonographers. We also explore the mechanisms that make systematic biases almost inevitable with the sample-based models.

SUBJECTS AND METHODS

Subjects

The data were collected over a period of 22 years, from 1988 to 2009. They comprise routine fetal ultrasound examinations performed in Oppland County, Norway, on an unselected population of pregnant women. There are two maternity wards in the county, at the Gjøvik and Lillehammer hospitals, each contributing from 800 to 1100 births every year. Most of the women in the study were examined prenatally and subsequently gave birth at these hospitals. In addition, data were collected from two smaller midwifery units at other locations in the same county.

Pregnancies with a fetal biparietal diameter (BPD) in the range of 38–60 mm or a femur length (FL) in the range of 21–42 mm at the routine ultrasound examination were included. Multiple pregnancies, pregnancies complicated by stillbirth, diagnosed anomalies, induction of labor for

reasons other than post-term pregnancy, or elective Cesarean sections were not included. In accordance with the post-term managing scheme of the maternity wards during the study period, inductions at eleven or more days past the EDD were defined as post-term inductions.

In total, fetal measurements from 23,020 second-trimester routine ultrasound examinations were included.

Prediction models

Three different models for estimating the date of delivery were evaluated in this study. The obstetric wheel 'Snurra' (referred to here as 'Trondheim–1984')[16] predicted GA from second-trimester BPD measurements between 38 and 60 mm only, and derived EDD by using a pregnancy duration of 282 days. The model was developed from ultrasound examinations of 90 carefully selected females with reliable menstrual data and anticipated normal pregnancies, included in a prospective, longitudinal study. A fourth-order polynomial regression analysis was used to establish the curves. This was the only dating method in use in Norway from 1984 until 2005.

The second sample-based method, 'Terminhjulet' ('Bergen–2004')[17, 18], used a similar statistical model based on fractional polynomial regression analysis[19], with a newer data sample. It predicts GA from either BPD (14–60 mm), FL (2–44 mm) or head circumference (50–134 mm) measurements, and derives EDD by using pregnancy duration of 282 days. The model was constructed from a cross-sectional study of 650 highly selected, healthy women with reliable menstrual data, assumed uncomplicated, singleton pregnancies, and with less than 14 days' disagreement between the LMP-based and the ultrasound-based EDD.

The third prediction model 'eSnurra' ('Trondheim–2007')[5] employs the new population-based approach with direct prediction of date of delivery. It is based on second-trimester fetal measurements from an unselected population of approximately 37,000 singleton pregnancies. The median remaining time of pregnancy was estimated, using a local linear quantile regression model. Trondheim–2007 predicts date of delivery from BPD (25–60 mm) or from FL (11–42 mm) measurements.

The application of each model's prediction table is described elsewhere[3, 4]. For the Trondheim–1984 and the Bergen–2004 models, the date of delivery is estimated by adding 282 days to the estimated LMP-date. For Trondheim–2007, predicted remaining time of pregnancy is found from the published tables[5].

Ultrasound examinations

The ultrasound examinations were for the most part performed by specially trained midwives at the hospitals or at the midwifery units, and the remaining (10–15% per year) by doctors at the hospitals or in private practice. There were 23 different examiners altogether. Four of the midwives each performed between 3200 and 7000 of the included ultrasound examinations. The data were prospectively registered in a database. A large proportion of the included data were measurements from pregnancy weeks 17–19. In general, all clinical problems were managed according to the EDD predicted by the Trondheim–1984-model.

The management of post-term pregnancies in the departments was modified during the study period. In the first years, induction of labor was scheduled around 14 days past the estimated date of delivery (≥296 days), while in later years the post-term inductions have gradually been

scheduled earlier — from 7–11 days past term.

The BPD and the FL were measured according to the standard method for fetal ultrasound biometry in Norway[3, 4]: BPD was measured from the outer to the outer contour of the parietal bones, and the mean of three BPD measurements was used. The FL was measured as the length of the ossified part of the femoral diaphysis in a longitudinal section, using the longest of three measurements[20]. The Bergen–2004 model used the mean of three FL measurements: this issue has been addressed previously[3].

Statistical methods

The three models for ultrasound-based prediction of date of delivery were applied to data collected from the ultrasound examinations and the subsequent deliveries. The 22,815 measurements with a BPD in the range 38–60 mm were used for EDD-calculations with all the three models. The 22,553 FL measurements between 21 and 42 mm were used with Bergen–2004 and Trondheim–2007.

To correct for the narrowed beam width in newer ultrasound scanners, which is demonstrated to shorten measurements in the lateral plane[21], a correction for the time-period[5] that applies to the FL measurements was included for the two newer models. The collection of the data in this study started in 1988, and newer prediction models should not be unrestrictedly applied to older data.

The resulting term predictions were compared with the actual time of delivery, and the disagreement was assessed in terms of the median bias for each model. The median bias reflects the systematic error of the term predictions and indicates the calibration level of the model as related to the study population. Predicting term too early results in a positive bias and an increased rate of apparently post-term pregnancies[3]. The median biases were calculated for subgroups with different fetal ages, in addition to the study population as a whole, since a varying bias may be missed if only the overall median bias is computed[22].

To assess the differences between the LMP-estimated GA at the actual time of the deliveries and the EDD as predicted from the BPD-measurements with each model, data from women with available LMP information were used. From 1999 onwards, the registry included information on whether the LMP information was certain or not; in this period only women with certain LMP data were included in the sub-analysis. As a result, 19,131 measurements were available. LMP was defined reliable when the woman was certain about the exact LMP date.

P-values for testing a non-zero median bias were computed using permutation tests with 2000 permutations. All analyses and graphics were produced in the R statistical programming environment[23].

RESULTS

Table 1 shows the percentage of ongoing pregnancies at 4, 7, 11 and 14 days past the EDD predicted from BPD and FL measurements by each model. Depending on the prediction model, there is a considerable difference in the percentage of pregnancies classified as post-term. This shows that the choice of dating model has a strong impact on post-term induction rates, regardless of which day past EDD that is recommended for post-term induction. As also indicated in the table, the rates of still ongoing pregnancies are nearly halved for all the models from day 7 to day 11 after EDD. The study population of 23 020 pregnancies, with large numbers even in the subgroups, resulted in only one median bias in one subgroup having non-significant *P*-value;

the bias of 0.13 days in the FL subgroup 21–26 mm for the Bergen–2004 model (P=0.18). All the other P-values were zero or < 0.01, indicating biases different from zero (results not shown). Hence, considering the large sample size, P-values were not very useful in deciding whether a bias was large enough to be clinically meaningful or not.

BPD-based predictions

Table 2 shows the median biases with 95% confidence intervals (CI) for the study group as a whole and for three different subgroups with BPD ranges corresponding to a GA of less than 18 weeks (38–43 mm), around 18 weeks (44–46 mm) and more than 18 weeks (47–60 mm). Figure 1 shows the median biases of the three models for each BPD value in the span of the study. The biases of the two sample-based models varied substantially, both within and between the models; the bias within a model was related to different BPD values. The bias for the population-based model was stable, essentially within ±1 day.

Figure 2 shows the GA at the actual time of the delivery as estimated from the LMP of the women with reliable LMP-data, compared with the EDDs predicted from BPD measurements with the three ultrasound models, in the same pregnancies. There is a consistently lower discrepancy between the EDD predictions from the population-based ultrasound model and the LMP-based GAs, than there is between the traditional ultrasound models' EDD predictions and the LMP-based GAs.

FL-based predictions

The two models, Bergen–2004 and Trondheim–2007, were applied to the FL measurements. Figure 3 shows the median biases of the two models for each FL value in the inclusion range. The bias of the sample-based model varied substantially with the size of the fetus at the time of the ultrasound examination, while it was essentially stable for the population-based model. The extent of the biases was similar to that observed with the BPD-based predictions for the same models, and also to the biases found for the same models when evaluated on another population[4].

The median biases with 95% CI for the study group as a whole and for three different subgroups with FL-ranges corresponding to a GA of less than 18 weeks (21–26 mm), around 18 weeks (27–29 mm) and above 18 weeks (30–42 mm) are shown in Table 3.

DISCUSSION

The present evaluation of a population-based model for prediction of date of delivery and the comparison of its predictions with those from two traditional models emphasizes the importance of continuous quality assessment. In this sample of 23 020 second-trimester examinations, the EDD-predictions from the BPD and FL measurements of the population-based model were reliable; the median bias was on the whole within ±1 day, confirming earlier findings[4]. The biases of the sample-based models were considerable in the study population as a whole and in the subgroups, both for BPD-based and FL-based predictions. The median biases varied substantially both with the fetal size at the time of the examination and between the two models, one generally predicting too early and the other too late (Figures 1 and 3). This also agrees with previous findings[3,4]. Both models performed adequately for a restricted span of fetal measurements. The EDD-discrepancy between them was consistently great, and never <2 days for the BPD-based predictions. For the late FL-based predictions

from Bergen–2004 the bias amounted to >4 days.

In 2006, data collected from the same study population during 1989–1999 were used to evaluate the two sample-based models' BPD-based predictions[24]. That study included only women with reliable LMP-dates; all inductions of labor were excluded, negatively affecting the mean bias[25]. An updated study was thus needed to evaluate the population-based model and the FL measurements using the more stable median bias as outcome.

To remove the overall median bias of the sample-based models one could add or subtract a constant value to all predictions[22]; this would correspond to shifting the curves in Figure 1 and 3 up or down along the y-axis until the median bias is zero. However, the slope of the curves would remain. Particularly for Bergen–2004, since both curves slope upward to the right, a correct overall calibration would result in EDD-predictions that are too late for the small fetuses and too early for the large ones. Thus, a simple calibration improving the overall bias would have unfortunate consequences. An optimal calibration should remove the bias over the whole range of measurement values; the population-based model achieves precisely that[4].

The population-based model was constructed from observations of the actual date of delivery to predict the remaining time of pregnancy and EDD from first- or second-trimester fetal measurements[5, 7]. However, modern pregnancy care requires a reliable EDD in the late stages of pregnancy, and knowledge of GA in the early stages. The traditional sample-based models were devised to estimate a hypothetical LMP, *i.e.,* GA, from second-trimester fetal measurements, and derive the EDD-prediction from this[3, 5, 6]. The population-based model estimates the GA as 283 minus the predicted remaining time of pregnancy. In the reference population[5, 26], 283 days is the median time from LMP to birth. Since the traditional models are based on estimating the LMP, one might assume that these methods would provide EDDs and GAs that correspond more closely to those computed from the LMP. Interestingly, this is not the case (Figure 2). The EDD-predictions from the population-based method correspond more closely to GA at delivery computed from the LMPs of women with reliable LMP-data, as seen from the narrower distribution curve of the discrepancy between the LMP- and ultrasound-based estimates. An overall calibration to remove these median differences (the shifts of the curves away from zero) would not alter the shape of the curves. Thus, the population-based predictions would still agree better with the LMP-based predictions.

The better corresponding ultrasound- and LMP-based estimates have immediate clinical consequences. First, it is beneficial both for scheduling examinations and reducing concern. Second, it reduces the risk of erroneous dating for fetuses that are small or large at the routine examination[27]. The new population-based method is thus better adapted to the actual target population than are the sample-based methods.

The present analysis, together with two previous studies[3, 4], comprises a total of 73 400 examinations in three different populations. These studies demonstrate that both sample-based models give systematically biased EDD-predictions and GA-calculations. The essential problem is that the traditional models were developed on samples with distributions different from the populations they are applied to. The population distribution has a strong concentration of examinations around 17–19 weeks. The considerable number of fetuses being small- and large-for-GA (SGA and

LGA) in this central group spill over to lower and higher BPD and FL values with fewer observations and pull their medians toward 17–18 weeks. These medians constitute the optimal predictions, paying attention to the average-for-GA (AGA) fetuses and to the spillover of the SGA/LGA fetuses. The sample-based models were developed on data with a flat GA-distribution, thus only paying attention to the AGA fetus. The SGA/LGA fetuses from more central dates will be erroneously interpreted as AGA fetuses and given a wrong age, producing the reported prediction biases. Conversely, since it is aimed at the actual population, the new method corrects for this effect.

The population-based model will better predict the date of delivery for fetuses with intrauterine growth restriction (IUGR). However, identifying early IUGR fetuses cannot be done from one single ultrasound examination, irrespective of prediction model[27]. Any significant difference between reliable LMP-based and ultrasound-based EDD-dates indicates a need for further evaluations[27-29].

In uncomplicated pregnancies with spontaneous deliveries close to the EDD, inaccurate dating is of less clinical importance, yet of interest in assessing perinatal outcome or evaluating management protocols. While the preterm deliveries mainly are unavoidable even if occasionally scheduled, iatrogenic post-maturity may follow a biased EDD, leading to unnecessary induction of labor shortly past term[4]. The resulting increase in wrongly identified post-term pregnancies is substantial (Table 2), yet often ignored. A prerequisite for comparison of induction routines is unbiased and uniform EDD-predictions with comparable post-term rates[12].

In conclusion, to obtain reliable EDD-predictions, the distribution of the population used to develop a model must correspond to the population to which the model is applied. The model must also answer dating questions both in early and late pregnancy. Including this sample of 23 020 examinations we now have confirmed our findings in a population totaling 73 400 examinations[3,4]. The population-based model is the method of choice for assessing GA and EDD.

ACKNOWLEDGEMENT

We would like to thank Nancy Lea Eik-Nes for revising the manuscript.

REFERENCES

1. Whitworth M, Bricker L, Neilson JP, Dowswell T. Ultrasound for fetal assessment in early pregnancy. *Cochrane Database Syst Rev* 2010. Issue **4;** DOI: 10.1002/14651858.
2. National Collaborating Centre for Women's and Children's Health. Antenatal care. Routine care for the healthy pregnant woman. National Institute for Health and Clinical Excellence: London, 2010; http://www.nice.org.uk/nicemedia/live/11947/40115/40115.pdf [Accessed 10 March 2011].
3. Økland I, Gjessing HK, Grøttum P, Eik-Nes SH. Biases of traditional term prediction models: results from different sample-based models evaluated on 41 343 ultrasound examinations. *Ultrasound Obstet Gynecol* 2010; **36:** 728-734.
4. Økland I, Gjessing HK, Grøttum P, Eggebø TM, Eik-Nes SH. A new population-based term prediction model vs. two traditional sample-based models: validation on 9046 ultrasound examinations. *Ultrasound Obstet Gynecol* 2011; **37:** 207-213.
5. 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.
6. Hutchon DJ, Ahmed F. Naegele's rule: a reappraisal. *Bjog* 2001; **108:** 775.
7. Salomon LJ, Pizzi C, Gasparrini A, Bernard JP, Ville Y. Prediction of the date of delivery based on first trimester ultrasound measurements: An independent method from estimated date of conception. *J Matern Fetal Neonatal Med* 2010; **23:** 1-9.

8. Hyett JA, Perdu M, Sharland GK, Snijders RS, Nicolaides KH. Increased nuchal translucency at 10-14 weeks of gestation as a marker for major cardiac defects. *Ultrasound Obstet Gynecol* 1997; **10:** 242-246.
9. Snijders RJ, Johnson S, Sebire NJ, Noble PL, Nicolaides KH. First-trimester ultrasound screening for chromosomal defects. *Ultrasound Obstet Gynecol* 1996; **7:** 216-226.
10. Miljeteig I, Markestad T, Norheim OF. Physicians' use of guidelines and attitudes to withholding and withdrawing treatment for extremely premature neonates in Norway. *Acta Paediatr* 2007; **96:** 825-829.
11. Pignotti MS. Extremely preterm births: recommendations for treatment in European countries. *Arch Dis Child Fetal Neonatal Ed* 2008; **93:** 403-406.
12. Zeitlin J, Blondel B, Alexander S, Breart G. Variation in rates of postterm birth in Europe: reality or artefact? *Bjog* 2007; **114:** 1097-1103.
13. Gulmezoglu AM, Crowther CA, Middleton P. Induction of labour for improving birth outcomes for women at or beyond term. *Cochrane Database Syst Rev* 2006. Issue **4;** DOI: 10.1002/14651858.
14. Lynch CD, Zhang J. The research implications of the selection of a gestational age estimation method. *Paediatr Perinat Epidemiol* 2007; **21 Suppl 2:** 86-96.
15. Blondel B, Morin I, Platt RW, Kramer MS, Usher R, Breart G. Algorithms for combining menstrual and ultrasound estimates of gestational age: consequences for rates of preterm and postterm birth. *Bjog* 2002; **109:** 718-720.
16. Eik-Nes SH, Grøttum P. Graviditetskalenderen Snurra. Scan-Med A/S, Drammen, Norway, 1983.
17. Johnsen SL, Rasmussen S, Sollien R, Kiserud T. Fetal age assessment based on ultrasound head biometry and the effect of maternal and fetal factors. *Acta Obstet Gynecol Scand* 2004; **83:** 716-723.
18. Johnsen SL, Rasmussen S, Sollien R, Kiserud T. Fetal age assessment based on femur length at 10-25 weeks of gestation, and reference ranges for femur length to head circumference ratios. *Acta Obstet Gynecol Scand* 2005; **84:** 725-733.
19. Altman DG, Chitty LS. New charts for ultrasound dating of pregnancy. *Ultrasound Obstet Gynecol* 1997; **10:** 174-191.
20. Hadlock FP, Harrist RB, Deter RL, Park SK. Fetal femur length as a predictor of menstrual age: sonographically measured. *AJR Am J Roentgenol* 1982; **138:** 875-878.
21. Økland I, Bjåstad TG, Johansen TF, Gjessing HK, Grøttum P, Eik-Nes SH. Narrowed beam width in newer ultrasound machines shortens measurements in the lateral plane: fetal measurement charts may be obsolete. *Ultrasound Obstet Gynecol* 2011; **38:** 82–86.
22. Gjessing HK, Grøttum P. Accuracy of second trimester fetal head circumference and biparietal diameter for predicting the time of spontaneous birth. *J Perinat Med* 2007; **35:** 350-351; author reply 351-352.
23. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, Austria, 2010; http://www.R-project.org. [Accessed 15 January 2011].
24. Backe B, Nakling J. Term prediction with ultrasound: evaluation of a new dating curve for biparietal diameter measurements. *Acta Obstet Gynecol Scand* 2006; **85:** 156-159.
25. Eik-Nes SH, Grøttum P, Gjessing HK. Regarding "Term prediction with ultrasound: evaluation of a new dating curve for biparietal diameter". *Acta Obstet Gynecol Scand* 2006; **85:** 1276-1278; author reply 1278-1279.
26. Tunon K, Eik-Nes SH, Grøttum P. A comparison between ultrasound and a reliable last menstrual period as predictors of the day of delivery in 15 000 examinations. *Ultrasound Obstet Gynecol* 1996; **8:** 178-185.
27. Verburg BO, Steegers EA, De Ridder M, Snijders RJ, Smith E, Hofman A, Moll HA, Jaddoe VW, Witteman JC. New charts for ultrasound dating of pregnancy and assessment of fetal growth: longitudinal data from a population-based cohort study. *Ultrasound Obstet Gynecol* 2008; **31:** 388-396.
28. Tunón K, Eik-Nes SH, Grøttum P. Fetal outcome when the ultrasound estimate of the day of delivery is more than 14 days later than the last menstrual period estimate. *Ultrasound Obstet Gynecol* 1999; **14:** 17-22.
29. Nguyen TH, Larsen T, Engholm G, Møller H. A discrepancy between gestational age estimated by last menstrual period and biparietal diameter may indicate an increased risk of fetal death and adverse pregnancy outcome. *Bjog* 2000; **107:** 1122-1129.

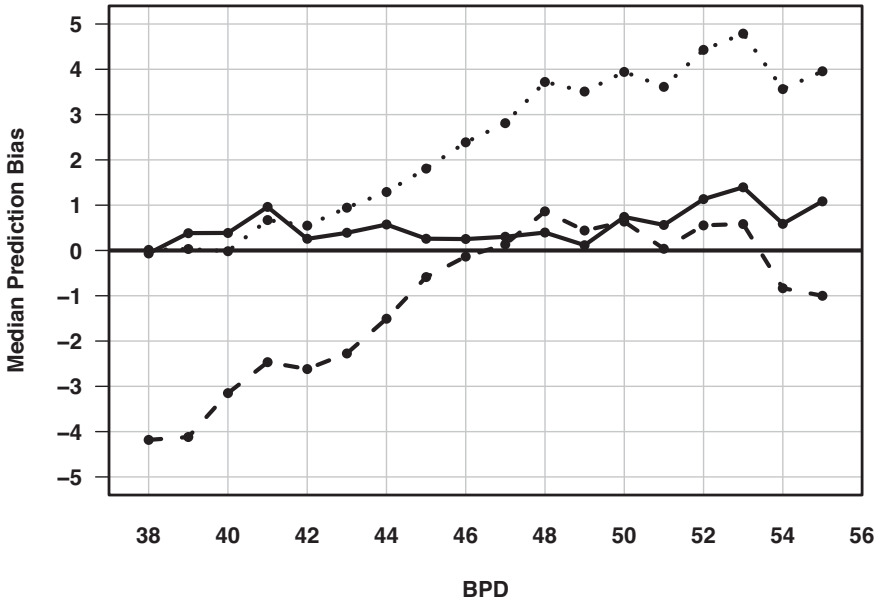


Figure 1 Median biases for the three models (Trondheim–1984 (dashed line), Bergen–2004 (dotted line) and Trondheim–2007 (solid line)) related to different biparietal diameter measurements.

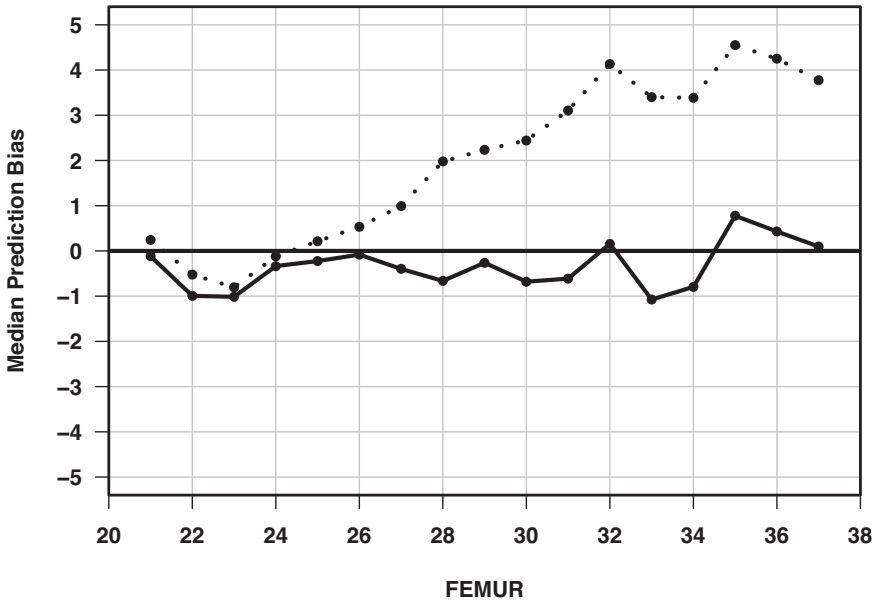


Figure 3 Median biases for the two models (Bergen–2004 (dotted line) and Trondheim–2007 (solid line)) related to different femur length measurements.

Table 1 The percentage of still ongoing pregnancies at 4, 7, 11 and 14 days past the date of delivery predicted by each model.

	BPD-based predictions			*FL-based predictions*	
Days past EDD	Trondheim–1984	Bergen–2004	Trondheim–2007	Bergen–2004	Trondheim–2007
4	30.7	43.6	36.6	41.5	32.9
7	19.9	30.6	24.4	28.4	21.4
11	9.2	16.6	11.9	15.2	9.9
14	4.0	9.2	5.8	7.8	4.4

Table 2 Term prediction from biparietal diameter (BPD) measurements with the three different models: Median bias for the three BPD-range groups and for the study group as a whole

	n	*Method*	*Median bias (days (95% CI))*
38 – 43	6074	Trondheim–1984	–2.75 (–3.01 to –2.48)
		Bergen–2004	0.22 (0.17 to 1.14)
		Trondheim–2007	0.38 (0.16 to 0.56)
44 – 46	8682	Trondheim–1984	–0.75 (–0.95 to –0.54)
		Bergen–2004	2.29 (1.52 to 2.29)
		Trondheim–2007	0.57 (0.26 to 0.73)
47 – 60	8059	Trondheim–1984	0.35 (0.11 to 0.58)
		Bergen–2004	3.68 (3.31 to 3.86)
		Trondheim–2007	0.40 (0.30 to 0.72)
38 – 60	22815	Trondheim–1984	–0.87 (–1.01 to –0.74)
		Bergen–2004	2.22 (2.14 to 2.29)
		Trondheim–2007	0.40 (0.30 to 0.57)

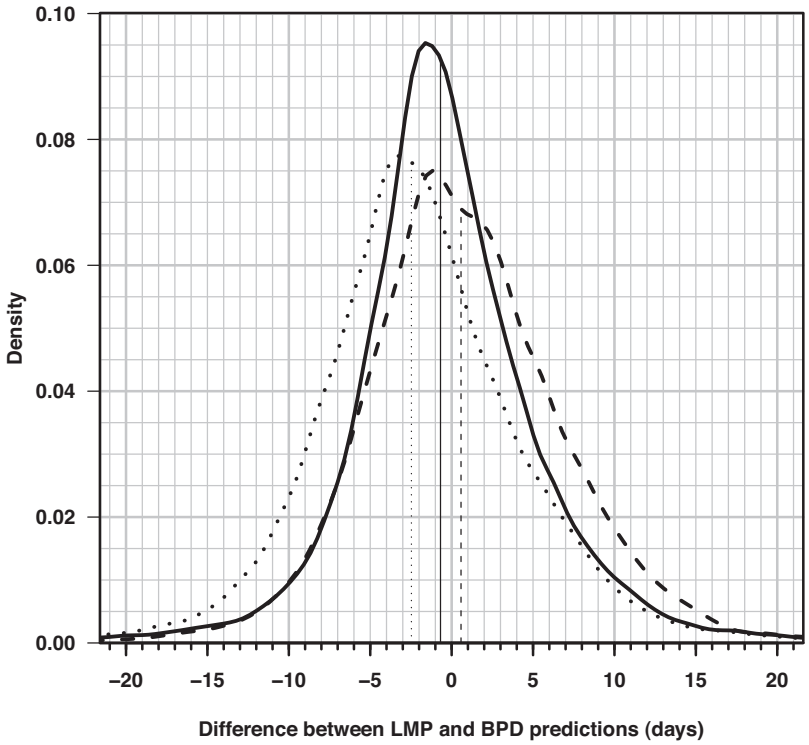


Figure 2 The difference between the term date computed from the last menstrual period and from biparietal diameter measurements with the three ultrasound models ((Trondheim–1984 (dashed line), Bergen–2004 (dotted line) and Trondheim–2007 (solid line)). The median difference is marked with vertical lines.

Table 3 Term prediction from femur length (FL) measurements with two different models: Median bias for the three FL-range groups and for the study group as a whole

FL (mm)	*n*	*Method*	*Median bias (days (95% CI))*
21 – 26	5753	Bergen–2004	0.13 (–0.10 to 0.29)
		Trondheim–2007	–0.27 (–0.74 to –0.08)
27 – 29	9315	Bergen–2004	1.72 (1.28 to 1.72)
		Trondheim–2007	–0.26 (–0.48 to –0.16)
30 – 42	7485	Bergen–2004	3.39 (2.98 to 3.61)
		Trondheim–2007	–0.46 (–0.63 to –0.41)
21 – 42	22553	Bergen–2004	1.72 (1.49 to 1.90)
		Trondheim–2007	–0.40 (–0.48 to –0.26)

Paper IV

Ultrasound Obstet Gynecol 2011; **38**: 82–87
Published online in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/uog.8954

Narrowed beam width in newer ultrasound machines shortens measurements in the lateral direction: fetal measurement charts may be obsolete

I. ØKLAND*†, T. G. BJÅSTAD‡, T. F. JOHANSEN‡, H. K. GJESSING§¶, P. GRØTTUM** and S. H. EIK-NES*††

**National Center for Fetal Medicine, Department of Obstetrics and Gynecology, St Olavs University Hospital, Trondheim, Norway; †Department of Obstetrics and Gynecology, Stavanger University Hospital, Stavanger, Norway; ‡Department of Circulation and Medical Imaging, Norwegian University of Science and Technology, Trondheim, Norway; §Division of Epidemiology, Norwegian Institute of Public Health, Oslo, Norway; ¶Department of Public Health and Primary Health Care, University of Bergen, Bergen, Norway; **Section of Medical Informatics, University of Oslo, Oslo, Norway; ††Department of Laboratory Medicine, Children's and Women's Health, Norwegian University of Science and Technology, Trondheim, Norway*

KEYWORDS: beam width; crown–rump length; date of delivery; femur length; fetal measurements; gestational age; resolution

ABSTRACT

Objectives *Fetal ultrasound measurements are made in axial, lateral and oblique directions. Lateral resolution is influenced by the beam width of the ultrasound system. To improve lateral resolution and image quality, the beam width has been made narrower; consequently, measurements in the lateral direction are affected and apparently made shorter, approaching the true length. The aims of this study were to explore our database to reveal time-dependent shortening of ultrasound measurements made in the lateral direction, and to assess the extent of beam-width changes by comparing beam-width measurements made on old and new ultrasound machines.*

Methods *A total of 41 941 femur length measurements, collected during the time-period 1987–2005, were analyzed, with time as a covariate. Using three ultrasound machines from the 1990s and three newer machines from 2007, we performed 25 series of blinded beam-width measurements on a tissue-mimicking phantom, measuring at depths of 3–8 cm with a 5-MHz transducer.*

Results *Regression analysis showed time to be a significant covariate. At the same gestational age, femur length measurement was 1.15 (95% CI, 1.08–1.23) mm shorter in the time-period 1999–2005 than in the time-period 1987–1992. Overall, the beam width was 1.08 (95% CI, 0.50–1.65) mm narrower with the new machines than with the old machines.*

Conclusions *Technical improvements in modern ultrasound machines that have reduced the beam width affect fetal measurements in the lateral direction. This has clinical implications and new measurement charts are needed.*

INTRODUCTION

Fetal biometry is an important part of obstetric ultrasound examination. An increasing number of fetal structures are being measured to assess fetal age and growth, risk of chromosomal aberrations and date of delivery[1,2]. As measurement deviations may lead to interventions, the correctness of each measurement is critical.

In two-dimensional imaging mode, fetal structures are measured in axial, lateral or varying oblique directions; in three-dimensional mode, the elevation plane may also be used. The measurement resolution in the various directions is dominated by differing physical features. The axial resolution is determined by the frequency and bandwidth of the transducer, and the lateral resolution is determined mainly by the beam width of the ultrasound system[3]. The axial resolution is superior to the lateral resolution[4].

In recent years, there have been several observations of apparently 'shorter' fetal parameters in newer measurement charts[5–9], the findings being attributed to improved machines and technical development in general, and to different populations and measuring techniques, as well as to unreliable pregnancy dating.

A reduced beam width has been essential in the development of modern ultrasound machines, resulting

Correspondence to: Dr I. Økland, Department of Obstetrics and Gynecology, Stavanger University Hospital, PO Box 8100, N-4068 Stavanger, Norway (e-mail: iok@lyse.net)

Accepted: 26 January 2011

in improved image resolution and better measurement quality[4]. Because of this narrowing in beam width, however, measurements of fetal structures in the lateral/oblique direction, such as femur length (FL), crown–rump length (CRL) and occipitofrontal diameter, have become shorter[10]. The beam-narrowing process causes the ultrasound-based measurements to approach the true length of the structure[10,11].

As fetal measurements are important predictors, potential sources of error should be analyzed[4]. Measurements of length must be reproducible, and they ought to be independent of scanner, transducer and measurement depth and orientation to be valid[12]. Being based on studies performed using ultrasound machines of a completely different generation, the use of old measurement charts may be in potential conflict with measurements from modern scanners from the last decade[10–12].

The aims of this study were to assess the significance of narrowing beam width over time, by exploring potential changes in laterally assessed measurements in a population-based database. In addition, we wanted to compare beam-width measurements in old and new scanners to evaluate the extent of the reduced beam width in modern ultrasound machines.

MATERIALS AND METHODS

Ultrasound examinations

The clinical measurements in this study originated from the second-trimester routine fetal ultrasound examinations at the National Center for Fetal Medicine at St Olavs University Hospital in Trondheim, Norway, and the data were prospectively registered in an electronic database over the period 1987–2005. In accordance with the Norwegian practice of routine ultrasound examinations, the majority of the data included were measurements from pregnancy weeks 17–19 (Table S1 online). The study population was non-selected and came from a geographically well-defined area. More than 30 experienced and formally trained midwives performed the ultrasound examinations. In this study, data from a total of 41 941 examinations in 38 725 pregnancies were included; for each, complete information about the date of the ultrasound scan and the date of delivery were available, and each had a fetal biparietal diameter (BPD) of 35–55 mm, a mean abdominal diameter (MAD) of 32–53 mm or a gestational age (GA) of 113–152 days (corresponding to 16 + 1 to 21 + 5 weeks) at the time of the ultrasound examination.

To assess the beam-width narrowing and its impact on measurements in the lateral direction, the data on the FL values from the ultrasound examinations were selected for evaluation.

Pregnancies with suspected manifold complications were not included in our data file[13]. However, pregnancies were included in which labor had been induced for post-term pregnancy[13]. The study material was divided into three, approximately equal, time-periods: 1987–1992 ($n = 13\,354$), 1993–1998 ($n = 13\,503$) and 1999–2005 ($n = 15\,084$). The GA was computed as described for the recently developed population-based term-prediction model[14].

The FL was measured with the femoral diaphysis in a longitudinal section, in accordance with the method described by Goldstein *et al.*[15]. At an angle of $< 45°$ to the horizontal plane, the ossified part of the diaphysis was measured three times on independently generated screen images, and the longest measurement was used, rounded to the nearest millimeter.

Ultrasound examinations were performed using Hitachi EUB-410, Hitachi EUB-415, Hitachi EUB-6000 and Hitachi EUB-6500 (Hitachi, Tokyo, Japan), Vingmed System Five (Vingmed Sound, Horten, Norway) and Logic 500 (GE Healthcare, Milwaukee, WI, USA) machines, all of which had curvilinear transducers with a frequency range of 3.5–5 MHz. From June 1988 to May 1989, 3.5- or 5-MHz transducers were used for the examinations. From June 1989 onwards, 5-MHz transducers were generally preferred[16].

Technical examinations

The ultrasound beam width was measured in three old and three new machines. The technically old machines were two Hitachi EUB-415 ultrasound scanners from 1993, and one Hitachi EUB-405 ultrasound scanner from 1996 (Hitachi, Tokyo, Japan), all with a 5-MHz probe of the type EUP C-324. The three new machines, which were produced in 2007, were a Siemens Acuson Antares (Siemens, Seattle, WA, USA) with a CH6-2 probe and two Hitachi HiVision 900E ultrasound scanners, both with a EUP-C524 probe; all of these probes operated at 5 MHz.

With each ultrasound machine, we carried out beam-width measurements on a tissue-mimicking phantom: CIRS model 40 (CIRS Tissue Simulation & Phantom Technology, Norfolk, VA, USA), measuring at six different depths, using the caliper function of the scanners. The phantom consists of several reflective strings at different depths, and enables measurements of the point-spread function (PSF) of an ultrasound imaging system, as seen in Figure S1 online. The width of the PSF relates directly to the resolution[17], and hence the beam width of the scanner, at a specific depth. Thus, the term 'beam width' refers to the width of the system's two-way beam, which is the effective beam resulting from the combination of the transmit and the receive focusing.

The focus was optimized for the imaging range of 3–8 cm. We were limited to two transmit focal zones with the old scanners and used the default of three transmit focal zones with the new ones. The screen images were zoomed to a depth of 3–8 cm before measuring.

As in clinical measuring, the beam width was defined at the middle gray tone level of the PSF image[3]. This level corresponds to the middle of the dynamic display range that is the range of echo intensities displayed on the screen. A simulated example of such a measurement is shown in Figure S1, where the referred width corresponds

to a 25-dB drop in echo intensity compared with that of the white level. To enable comparison of beam widths among different machines, the gain and dynamic display range were adjusted to be as similar as possible in old and new machines. The gain was adjusted until an equally bright perceived background speckle pattern at all depths of the phantom was obtained. The dynamic display range in the new scanners was reduced in order to resemble the dynamic display range of the old ones.

The probe was positioned perpendicular to the phantom surface and fixed using a mechanical arm. One operator performed all measurements on the phantom. The sequence of six measurements (at depths 3, 4, 5, 6, 7 and 8 cm) was repeated 25 times with each scanner. The operator was blinded to the beam-width measurements during the procedure. Between each series, the phantom was shifted slightly in the lateral direction to obtain a minor alteration in the ultrasound image. Thus, we eliminated the effect of caliper inaccuracy and included the possible effect of shift variance, that is, a spatial increase and variation in the PSF width caused by a too low scan-line density.

Statistical methods

We analyzed the 41 941 FL measurements to evaluate the possible effect of a changing beam width over time. First, median FL values were computed for each day of GA in each of the three time-periods. Similarly, median FL values were computed for each BPD (in mm) and each MAD (in mm) in the three time-periods. This allowed an assessment of whether median FL values vary over time, independently of fetal size and age. We thus controlled for potential changes in fetal growth pattern or in the time of routine ultrasound examinations. As FL measurements are given as integer values we computed the median as a linear interpolation between the two closest integer values. Second, to obtain a summary of the change in median FL values over the three time-periods, we analyzed the data using a quantile regression model, with FL as the dependent variable and time-period as the categorical variable. Three separate analyses were carried out, adjusting for GA, BPD and MAD, respectively. As the effects of GA, BPD and MAD on FL are very close to linear in the relevant range, the variables were included as linear covariates in the models. Third, to obtain a more detailed picture of the change in FL over time, we performed the same quantile regression analyses, replacing the three time-period categories with narrower time-period categories spanning one year each (i.e. every year from 1987 to 2005 as a single-year category). Again adjusting linearly for GA, BPD and MAD, we obtained median FL values for each 1-year category, standardized for GA, BPD and MAD separately.

To analyze the beam-width measurements from the phantom we used a linear mixed-effects regression model. We regressed the measured beam width on the measurement depth and machine generation. Measurement depth had a close to linear effect on beam width and was thus included as a linear covariate. Machine generation was used as a covariate with two levels: old and new. As 25 measurement replications were made for each depth on each of the six ultrasound machines, we controlled for within-machine dependent measurements by adding machine to the model as a random effect with six levels.

All analyses and graphs were produced in the R statistical programming environment[18]. Quantile regression was carried out with the *rq* function in the *quantreg* library[19], and the linear mixed-effects model was estimated using *lme* from the *nlme* library[20].

RESULTS

The regression analyses showed that time was a significant covariate. At the same GA, the median FL measured 1.15 (95% CI, 1.08–1.23) mm shorter ($P < 0.005$) in the third time-period (1999–2005) than in the first (1987–1992), as shown in Figure S2 online. As the curves indicate, there was a smaller difference, of only −0.16 mm ($P < 0.005$), between the first and the second (1993–1998) time-periods.

The regressions of the FL measurements on the BPD measurements demonstrated that for the same BPD, the median FL was also significantly shorter; 0.98 (95% CI, 0.93–1.04) mm ($P < 0.005$) shorter in the third time-period than in the first, as shown in Figure S3 online. Between the first and the second time-periods, the change in FL measurements was −0.37 mm ($P < 0.005$).

Regressions of the FL measurements on the MAD measurements showed corresponding shifts, but on a smaller scale: a minor shortening of 0.59 (95% CI, 0.54–0.63) mm ($P < 0.005$) from the first to the third time-period, and an insignificant median FL difference of −0.01 mm from the first to the second time-period.

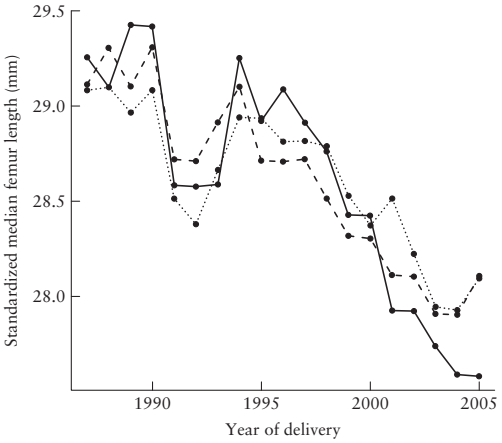


Figure 1 Changes in median femur length for every 1-year period between 1987 and 2005, adjusted linearly and standardized for gestational age (——), biparietal diameter (- - - - -) and mean abdominal diameter (. . . . .).

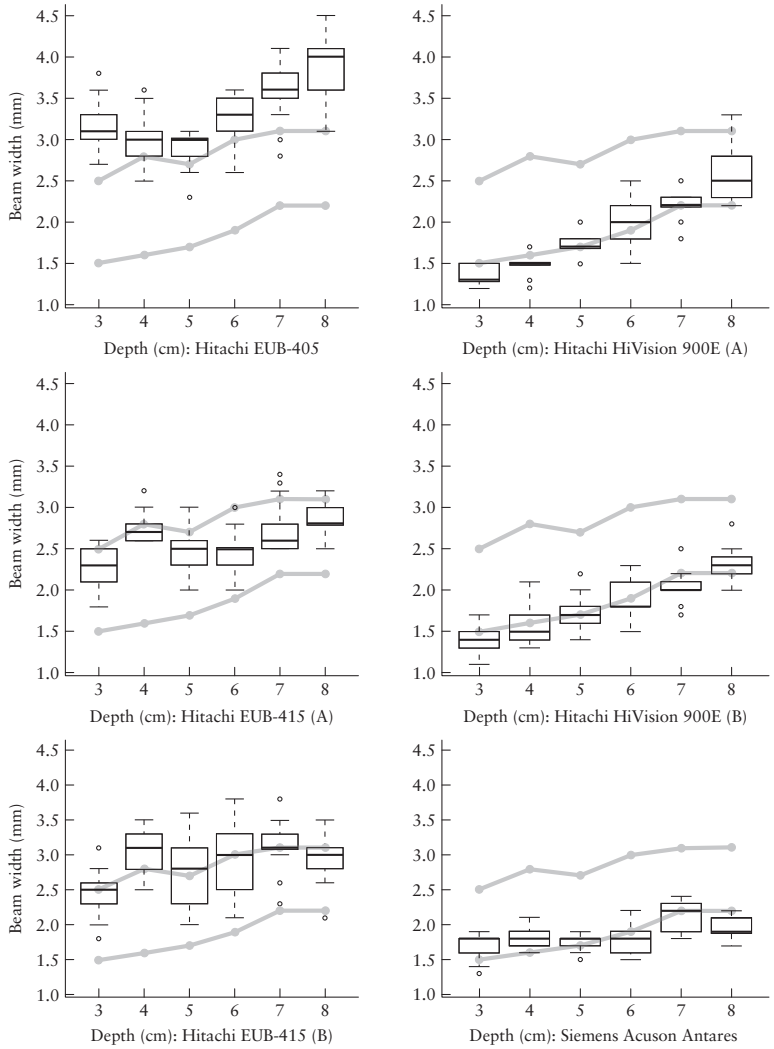


Figure 2 Box-plots of beam-width measurements for each of the six ultrasound machines. New machines are presented on the right. The gray background curves represent the median beam widths at the various measurement depths for the new (lower curve) and the old (upper curve) scanners. Median, interquartile range, range and outliers are shown.

Figure 1 shows the estimation of the changes in median FL for each year of the study period, adjusted linearly and standardized for GA, BPD and MAD, respectively. The figure illustrates the main trend and reflects the periods of major replacement of ultrasound machines in our department.

For the beam-width measurements on the phantom, the overall median beam width was 1.08 (95% CI, 0.50–1.65) mm narrower with the new machines than with the old machines ($P = 0.006$). For both scanner generations the measurements increased significantly with increasing depths. Figure 2 shows the results, as box-plots, for each of the six machines.

DISCUSSION

In this study we explored our database for time-dependent changes in measurements made in lateral directions and found that FL measured approximately 1 mm shorter at the same GA in the third time-period than in the first. We therefore carried out beam-width measurements on a tissue-mimicking phantom and compared the measurements from old and new ultrasound scanners. The beam width was narrower with the new machines. Technical improvements, particularly those which have reduced beam width in modern ultrasound machines, affect certain ultrasound measurements[10], and

the non-axial measurements are typically influenced the most. This effect may have clinical implications for ultrasound-based fetal age assessment.

In 1975, Robinson and Fleming published a critical evaluation[11] of a 2-year-old preliminary communication on sonar CRL measurements[21]. The beam-width effect was given particular focus in the analysis of random and systematic measurement errors, resulting in a 'corrected' CRL table[11]. The corrected curve had systematically shorter CRL values, resulting in a difference of 1–2 days of calculated GA for a given CRL[22]. Interestingly, the existence of two different Robinson and Fleming reference curves is not well known. The beam-width problem was hardly raised again until 20 years later, when Jago *et al.*, in a small-scale study, found that the FL was significantly longer when measured with an old scanner than with a new one[10]. They indicated the effect of beam-width narrowing over time and its potential consequences for measurements in the lateral direction. In a population-based database with 41 941 clinical measurements, our results confirmed their suggestion. Moreover, our alternative beam-width measurement approach, which included more than two scanners and the effect of large scan-line spacing, also confirmed their findings.

A potential variation in a certain fetal measurement, for instance FL or BPD for the same GA, might be regarded as resulting from a possible alteration in fetal growth pattern over time. On the contrary, the relationship between the various parameters (e.g. how FL relates to BPD/MAD over time) indicates how the fetal geometry is being influenced, in other words, how the measurements vary over time, independently of fetal size and age.

Fetal measurements are made in axial, lateral and oblique directions. The BPD is a standard axial measurement, while the occipitofrontal diameter is measured laterally, and FL, CRL and many other biometric structures are measured more or less obliquely[23]. In our database, FL, BPD and MAD measurements were available for analysis. Circumference measurements (and MAD) are, in essence, an approximately equal combination of an axial and a lateral diameter measurement, and are thus correspondingly affected by the reduced beam width. Therefore, the FL was shortened less when related to MAD measurements than to BPD measurements (Figure 1).

Echoes originating from the full width of an ultrasound beam will be displayed along the centerline of the beam. Consequently, in the same way as the single point in Figure S1, the femur will be extended laterally on an ultrasound screen image. The addition of a fixed beam-width dimension will affect shorter structures relatively more than longer structures. When imaging a nearly horizontally oriented femur, the echoes from each femoral end are 'picked up' by beams whose centerline is outside the actual ends. For all horizontally measured structures the image extends sideways up to 0.5 beam width at each end – one beam width in total – in excess of the true measurement, as shown schematically in Figure 3. For a femur at 45° the overestimation is reduced to one beam width multiplied by cosine 45°; 0.7 beam width in total, that is 0.35 at each end[10]. Therefore, one should measure the FL diagonally rather than horizontally, in order to reduce the negative beam-width effect.

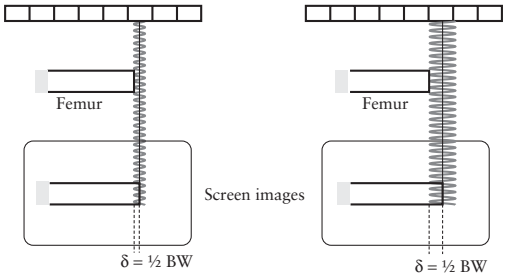


Figure 3 Schematic illustration of the narrow beam-width (BW) in new machines and the wide BW in old machines, showing how the width of the ultrasound beam influences the lateral extension of an object.

Over the years, various technical improvements have resulted in narrower ultrasound beams. The principal reason for obtaining a shorter FL measurement in the third time-period compared with the first and second time-periods (Figure 1) was probably a result of the transition from analog to digital beam-formation hardware. In the middle/late 1990s, digital processing hardware (such as analog to digital converters, application-specific integrated circuits (ASICs) and digital signal processors) was included in the scanners. This allowed for much more flexible and precise control of the beam-formation process, resulting in a narrower beam width throughout the whole ultrasound imaging sector[24–26]. Additionally, an increased system channel number enabled the use of larger apertures, generating a narrower beam width. The increased processing power also allowed a high scan-line density. Spacing the scan lines too far apart decreases the lateral resolution by more than that determined by the beam width alone.

Fetal measurements are essential for assessing fetal age[11,14] and growth[8,27], and in fetal anomaly scanning[2,7]. The measured parameter is related to a measurement chart, optimized for the purpose of the examination. Curves from the 1980s and 1990s are still widely used as reference charts[28], despite the fact that charts from the last decade typically are constructed from significantly shorter FL and CRL measurements[6–8]. The lack of knowledge of technical development and its consequences for ultrasound equipment in clinical use cause systematic errors, particularly in first-trimester assessment of GA, giving dating errors of up to half a week[4,6] and erroneous risk estimations in prenatal screening[22]. With the use of a modern scanner and an old dating chart, the FL/CRL will be considered 'too short'[10]; fetal age is then underestimated and the predicted date of delivery is set too late, which of course may have clinical consequences[29]. Once again, this underlines the importance of continuous quality assurance[30] and general guidelines[23] to standardize fetal biometric assessment.

In conclusion, measurements on a phantom confirmed the narrowing of beam width in newer ultrasound machines. Over time, measurements in the lateral direction have been affected, as verified by the shortening of FL measurements over an 18-year period in a large population. The lack of standardization in determining fetal age is challenging and has significant consequences, as clinical management may depend on a CRL/FL-based fetal age. Older reference charts have become obsolete, and new curves from recent studies using modern scanners are preferable.

ACKNOWLEDGMENT

We would like to thank Nancy Lea Eik-Nes for revising the manuscript.

REFERENCES

1. Degani S. Fetal biometry: clinical, pathological, and technical considerations. *Obstet Gynecol Surv* 2001; **56**: 159–167.
2. Snijders RJ, Nicolaides KH. Fetal biometry at 14–40 weeks' gestation. *Ultrasound Obstet Gynecol* 1994; **4**: 34–48.
3. DuBose TJ, Hill LW. Physics and methodology of obstetrical sonography. In *Fetal Sonography*, DuBose TJ (ed). W.B. Saunders Company: Philadelphia, PA, 1996; 73–75.
4. Pretorius DH, Nelson TR, Manco-Johnson ML. Fetal age estimation by ultrasound: the impact of measurement errors. *Radiology* 1984; **152**: 763–766.
5. MacGregor SN, Tamura RK, Sabbagha RE, Minogue JP, Gibson ME, Hoffman DI. Underestimation of gestational age by conventional crown-rump length dating curves. *Obstet Gynecol* 1987; **70**: 344–348.
6. Pexsters A, Daemen A, Bottomley C, Van Schoubroeck D, De Catte L, De Moor B, D'Hooghe T, Lees C, Timmerman D, Bourne T. New crown-rump length curve based on over 3500 pregnancies. *Ultrasound Obstet Gynecol* 2010; **35**: 650–655.
7. Longo D, DeFigueiredo D, Cicero S, Sacchini C, Nicolaides KH. Femur and humerus length in trisomy 21 fetuses at 11–14 weeks of gestation. *Ultrasound Obstet Gynecol* 2004; **23**: 143–147.
8. Verburg BO, Steegers EA, De Ridder M, Snijders RJ, Smith E, Hofman A, Moll HA, Jaddoe VW, Witteman JC. New charts for ultrasound dating of pregnancy and assessment of fetal growth: longitudinal data from a population-based cohort study. *Ultrasound Obstet Gynecol* 2008; **31**: 388–396.
9. Leung TN, Pang MW, Daljit SS, Leung TY, Poon CF, Wong SM, Lau TK. Fetal biometry in ethnic Chinese: biparietal diameter, head circumference, abdominal circumference and femur length. *Ultrasound Obstet Gynecol* 2008; **31**: 321–327.
10. Jago JR, Whittingham TA, Heslop R. The influence of ultrasound scanner beam width on femur length measurements. *Ultrasound Med Biol* 1994; **20**: 699–703.
11. Robinson HP, Fleming JE. A critical evaluation of sonar 'crown-rump length' measurements. *Br J Obstet Gynaecol* 1975; **82**: 702–710.
12. Gamba JL, Bowie JD, Dodson WC, Hedlund LW. Accuracy of ultrasound in fetal femur length determination. Ultrasound phantom study. *Invest Radiol* 1985; **20**: 316–323.
13. Økland I, Gjessing HK, Grøttum P, Eik-Nes SH. Biases of traditional term prediction models: results from different sample-based models evaluated on 41 343 ultrasound examinations. *Ultrasound Obstet Gynecol* 2010; **36**: 728–734.
14. 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.
15. Goldstein RB, Filly RA, Simpson G. Pitfalls in femur length measurements. *J Ultrasound Med* 1987; **6**: 203–207.
16. Tegnander E, Eik-Nes SH, Linker DT. Incorporating the four-chamber view of the fetal heart into the second-trimester routine fetal examination. *Ultrasound Obstet Gynecol* 1994; **4**: 24–28.
17. Szabo TL. Array beamforming. In *Diagnostic Ultrasound Imaging: Inside Out*, Szabo TL (ed). Elsevier Academic Press: London, 2004; 190–196.
18. R Development Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing: Vienna, 2010; http://www.R-project.org [Accessed 17 June 2010].
19. Koenker R. quantreg: Quantile Regression. R package version 4.44: Vienna, 2009; http://cran.r-project.org/web/packages/quantreg/index.html [Accessed 15 October 2010].
20. Pinheiro J, Bates D, DebRoy S, Sarkar D, R Development Core Team. nlme: Linear and nonlinear mixed effects models. R package version 3.1–96: Vienna, 2009; http://www.R-project.org [Accessed 15 October 2010].
21. Robinson HP. Sonar measurement of fetal crown-rump length as means of assessing maturity in first trimester of pregnancy. *Br Med J* 1973; **4**: 28–31.
22. Koster MP, Van Leeuwen-Spruijt M, Wortelboer EJ, Stoutenbeek P, Elvers LH, Loeber JG, Visser GH, Schielen PC. Lack of standardization in determining gestational age for prenatal screening. *Ultrasound Obstet Gynecol* 2008; **32**: 607–611.
23. Salomon LJ, Alfirevic Z, Berghella V, Bilardo C, Hernandez-Andrade E, Johnsen SL, Kalache K, Leung KY, Malinger G, Munoz H, Prefumo F, Toi A, Lee W; ISUOG Clinical Standards Committee. Practice guidelines for performance of the routine mid-trimester fetal ultrasound scan. *Ultrasound Obstet Gynecol* 2011; **37**: 116–126.
24. Song TK, Park SB. A new digital phased array system for dynamic focusing and steering with reduced sampling rate. *Ultrason Imaging* 1990; **12**: 1–16.
25. Steinberg BD. Digital beamforming in ultrasound. *IEEE Trans Ultrason Ferroelectr Freq Control* 1992; **39**: 716–721.
26. Holm S. Medical ultrasound transducers and beamforming. In *Proceedings of the 15th International Congress on Acoustics*. Trondheim, Norway, June 1995; 339–342.
27. Sahota DS, Kagan KO, Lau TK, Leung TY, Nicolaides KH. Customized birth weight: coefficients and validation of models in a UK population. *Ultrasound Obstet Gynecol* 2008; **32**: 884–889.
28. Saltvedt S, Almstrom H, Kublickas M, Reilly M, Valentin L, Grunewald C. Ultrasound dating at 12–14 or 15–20 weeks of gestation? A prospective cross-validation of established dating formulae in a population of in-vitro fertilized pregnancies randomized to early or late dating scan. *Ultrasound Obstet Gynecol* 2004; **24**: 42–50.
29. Økland I, Gjessing HK, Grøttum P, Eggebø TM, Eik-Nes SH. A new population-based term prediction model vs. two traditional sample-based models: validation on 9046 ultrasound examinations. *Ultrasound Obstet Gynecol* 2011; **37**: 207–213.
30. Ville Y. 'Ceci n'est pas une Èchographie': a plea for quality assessment in prenatal ultrasound. *Ultrasound Obstet Gynecol* 2008; **31**: 1–5.

SUPPORTING INFORMATION ON THE INTERNET

Table S1 and Figures S1–S3 may be found in the online version of this article.

SUPPORTING INFORMATION ON THE INTERNET

Table S1 and Figures S1–S3 may be found in the online version of this article.

Table S1 The distribution of examinations over the days of gestational age included in the study, corresponding to the range of 113 days (16 + 1 weeks) to 152 days (21 + 5 weeks).

Gestational age (days)	*Examinations (n)*
113–116	1730
117–120	3012
121–124	4516
125–128	5818
129–132	6091
133–136	5671
137–140	4316
141–144	3179
145–148	2101
149–152	1496
Total	37 930

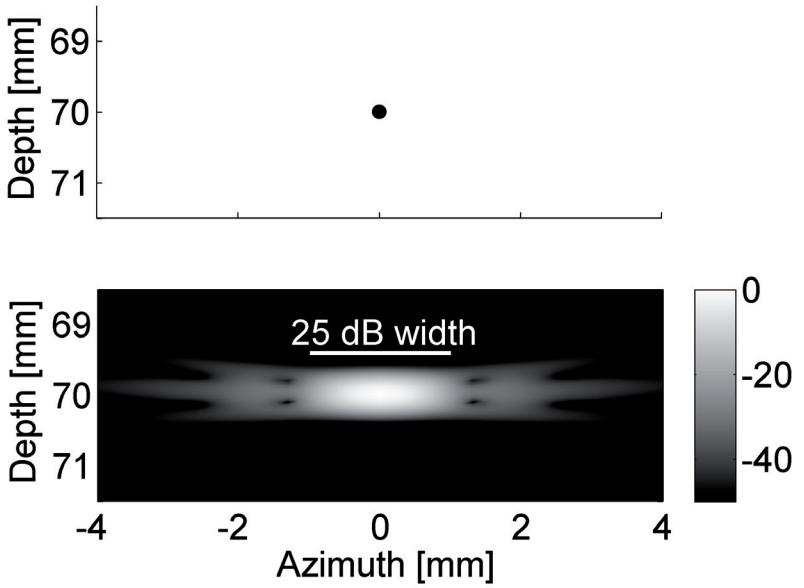


Figure S1 Ultrasound image of an infinitely small point also referred to as the point spread function (PSF) of the ultrasound imaging system. (a) The point placed in a diagram. (b) Simulated ultrasound screen image of the single point. The 25 dB width of the PSF is indicated.

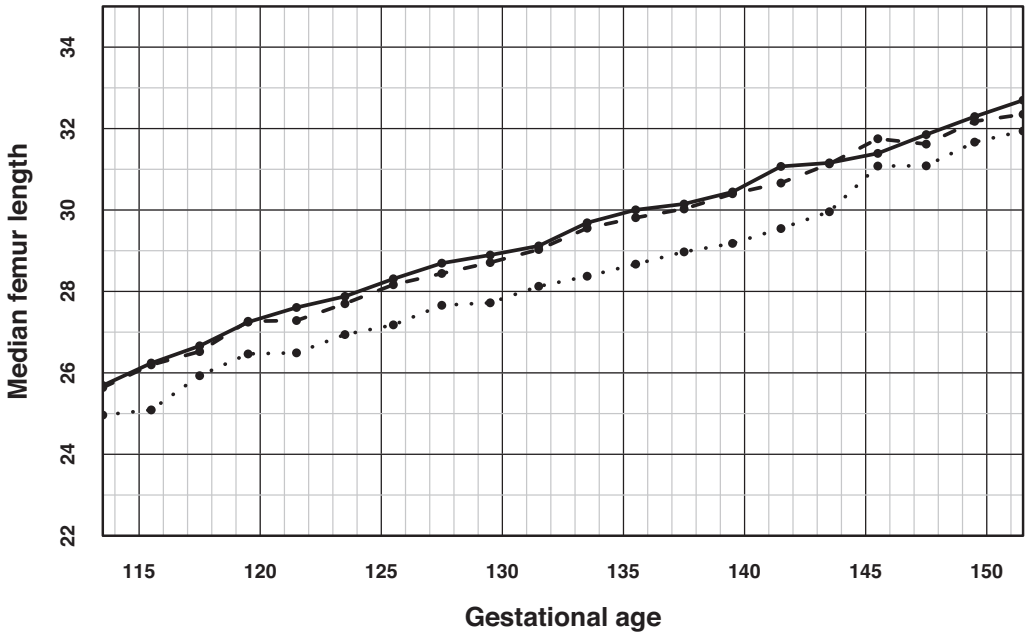


Figure S2 The median femur length measurements in the three different time periods related to days of gestational age. 1987–92 (unbroken line), 1993–98 (broken line), 1999–2005 (dotted line).

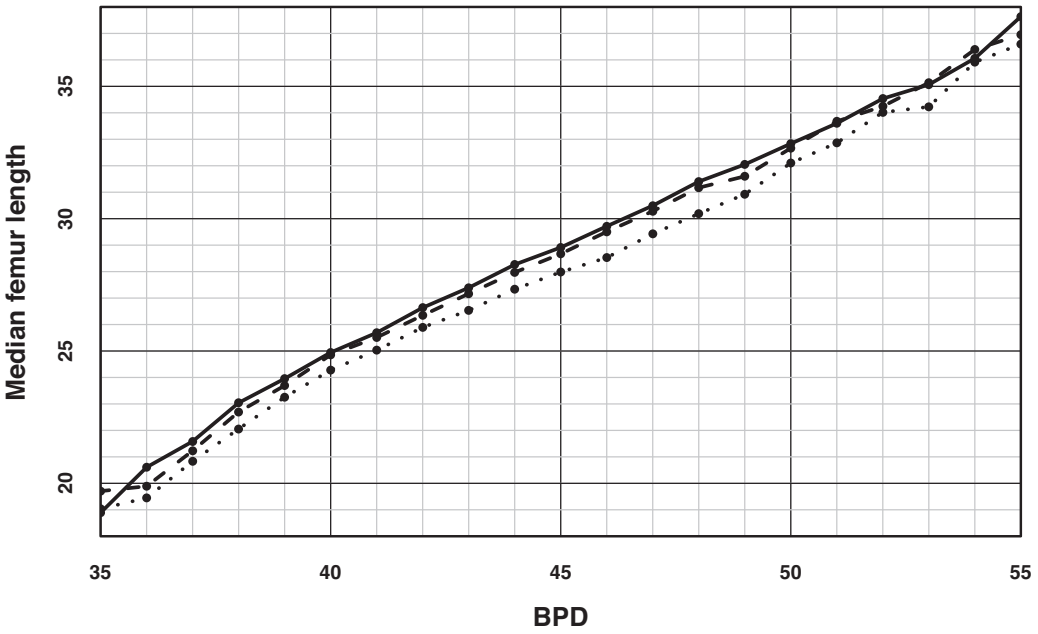


Figure S3 The median femur length measurements in the three time periods related to different measurements of biparietal diameter. 1987–92 (unbroken line), 1993–98 (broken line), 1999–2005 (dotted line).

Dissertations at the Faculty of Medicine, NTNU

1977

1. Knut Joachim Berg: EFFECT OF ACETYLSALICYLIC ACID ON RENAL FUNCTION
2. Karl Erik Viken and Arne Ødegaard: STUDIES ON HUMAN MONOCYTES CULTURED *IN VITRO*

1978

3. Karel Bjørn Cyvin: CONGENITAL DISLOCATION OF THE HIP JOINT.
4. Alf O. Brubakk: METHODS FOR STUDYING FLOW DYNAMICS IN THE LEFT VENTRICLE AND THE AORTA IN MAN.

1979

5. Geirmund Unsgaard: CYTOSTATIC AND IMMUNOREGULATORY ABILITIES OF HUMAN BLOOD MONOCYTES CULTURED IN VITRO

1980

6. Størker Jørstad: URAEMIC TOXINS
7. Arne Olav Jenssen: SOME RHEOLOGICAL, CHEMICAL AND STRUCTURAL PROPERTIES OF MUCOID SPUTUM FROM PATIENTS WITH CHRONIC OBSTRUCTIVE BRONCHITIS

1981

8. Jens Hammerstrøm: CYTOSTATIC AND CYTOLYTIC ACTIVITY OF HUMAN MONOCYTES AND EFFUSION MACROPHAGES AGAINST TUMOR CELLS *IN VITRO*

1983

9. Tore Syversen: EFFECTS OF METHYLMERCURY ON RAT BRAIN PROTEIN.
10. Torbjørn Iversen: SQUAMOUS CELL CARCINOMA OF THE VULVA.

1984

11. Tor-Erik Widerøe: ASPECTS OF CONTINUOUS AMBULATORY PERITONEAL DIALYSIS.
12. Anton Hole: ALTERATIONS OF MONOCYTE AND LYMPHOCYTE FUNCTIONS IN REALTION TO SURGERY UNDER EPIDURAL OR GENERAL ANAESTHESIA.
13. Terje Terjesen: FRACTURE HEALING AND STRESS-PROTECTION AFTER METAL PLATE FIXATION AND EXTERNAL FIXATION.
14. Carsten Saunte: CLUSTER HEADACHE SYNDROME.
15. Inggard Lereim: TRAFFIC ACCIDENTS AND THEIR CONSEQUENCES.
16. Bjørn Magne Eggen: STUDIES IN CYTOTOXICITY IN HUMAN ADHERENT MONONUCLEAR BLOOD CELLS.
17. Trond Haug: FACTORS REGULATING BEHAVIORAL EFFECTS OG DRUGS.

1985

18. Sven Erik Gisvold: RESUSCITATION AFTER COMPLETE GLOBAL BRAIN ISCHEMIA.
19. Terje Espevik: THE CYTOSKELETON OF HUMAN MONOCYTES.
20. Lars Bevanger: STUDIES OF THE Ibc (c) PROTEIN ANTIGENS OF GROUP B STREPTOCOCCI.
21. Ole-Jan Iversen: RETROVIRUS-LIKE PARTICLES IN THE PATHOGENESIS OF PSORIASIS.
22. Lasse Eriksen: EVALUATION AND TREATMENT OF ALCOHOL DEPENDENT BEHAVIOUR.
23. Per I. Lundmo: ANDROGEN METABOLISM IN THE PROSTATE.

1986

24. Dagfinn Berntzen: ANALYSIS AND MANAGEMENT OF EXPERIMENTAL AND CLINICAL PAIN.
25. Odd Arnold Kildahl-Andersen: PRODUCTION AND CHARACTERIZATION OF MONOCYTE-DERIVED CYTOTOXIN AND ITS ROLE IN MONOCYTE-MEDIATED CYTOTOXICITY.
26. Ola Dale: VOLATILE ANAESTHETICS.

1987

27. Per Martin Kleveland: STUDIES ON GASTRIN.
28. Audun N. Øksendal: THE CALCIUM PARADOX AND THE HEART.
29. Vilhjalmur R. Finsen: HIP FRACTURES

1988

30. Rigmor Austgulen: TUMOR NECROSIS FACTOR: A MONOCYTE-DERIVED REGULATOR OF CELLULAR GROWTH.
31. Tom-Harald Edna: HEAD INJURIES ADMITTED TO HOSPITAL.
32. Joseph D. Borsi: NEW ASPECTS OF THE CLINICAL PHARMACOKINETICS OF METHOTREXATE.
33. Olav F. M. Sellevold: GLUCOCORTICOIDS IN MYOCARDIAL PROTECTION.
34. Terje Skjærpe: NONINVASIVE QUANTITATION OF GLOBAL PARAMETERS ON LEFT VENTRICULAR FUNCTION: THE SYSTOLIC PULMONARY ARTERY PRESSURE AND CARDIAC OUTPUT.
35. Eyvind Rødahl: STUDIES OF IMMUNE COMPLEXES AND RETROVIRUS-LIKE ANTIGENS IN PATIENTS WITH ANKYLOSING SPONDYLITIS.
36. Ketil Thorstensen: STUDIES ON THE MECHANISMS OF CELLULAR UPTAKE OF IRON FROM TRANSFERRIN.
37. Anna Midelfart: STUDIES OF THE MECHANISMS OF ION AND FLUID TRANSPORT IN THE BOVINE CORNEA.
38. Eirik Helseth: GROWTH AND PLASMINOGEN ACTIVATOR ACTIVITY OF HUMAN GLIOMAS AND BRAIN METASTASES - WITH SPECIAL REFERENCE TO TRANSFORMING GROWTH FACTOR BETA AND THE EPIDERMAL GROWTH FACTOR RECEPTOR.
39. Petter C. Borchgrevink: MAGNESIUM AND THE ISCHEMIC HEART.
40. Kjell-Arne Rein: THE EFFECT OF EXTRACORPOREAL CIRCULATION ON SUBCUTANEOUS TRANSCAPILLARY FLUID BALANCE.
41. Arne Kristian Sandvik: RAT GASTRIC HISTAMINE.
42. Carl Bredo Dahl: ANIMAL MODELS IN PSYCHIATRY.

1989

43. Torbjørn A. Fredriksen: CERVICOGENIC HEADACHE.
44. Rolf A. Walstad: CEFTAZIDIME.
45. Rolf Salvesen: THE PUPIL IN CLUSTER HEADACHE.
46. Nils Petter Jørgensen: DRUG EXPOSURE IN EARLY PREGNANCY.
47. Johan C. Ræder: PREMEDICATION AND GENERAL ANAESTHESIA IN OUTPATIENT GYNECOLOGICAL SURGERY.
48. M. R. Shalaby: IMMUNOREGULATORY PROPERTIES OF TNF-α AND THE RELATED CYTOKINES.
49. Anders Waage: THE COMPLEX PATTERN OF CYTOKINES IN SEPTIC SHOCK.
50. Bjarne Christian Eriksen: ELECTROSTIMULATION OF THE PELVIC FLOOR IN FEMALE URINARY INCONTINENCE.
51. Tore B. Halvorsen: PROGNOSTIC FACTORS IN COLORECTAL CANCER.

1990

52. Asbjørn Nordby: CELLULAR TOXICITY OF ROENTGEN CONTRAST MEDIA.
53. Kåre E. Tvedt: X-RAY MICROANALYSIS OF BIOLOGICAL MATERIAL.
54. Tore C. Stiles: COGNITIVE VULNERABILITY FACTORS IN THE DEVELOPMENT AND MAINTENANCE OF DEPRESSION.
55. Eva Hofsli: TUMOR NECROSIS FACTOR AND MULTIDRUG RESISTANCE.
56. Helge S. Haarstad: TROPHIC EFFECTS OF CHOLECYSTOKININ AND SECRETIN ON THE RAT PANCREAS.
57. Lars Engebretsen: TREATMENT OF ACUTE ANTERIOR CRUCIATE LIGAMENT INJURIES.
58. Tarjei Rygnestad: DELIBERATE SELF-POISONING IN TRONDHEIM.
59. Arne Z. Henriksen: STUDIES ON CONSERVED ANTIGENIC DOMAINS ON MAJOR OUTER MEMBRANE PROTEINS FROM ENTEROBACTERIA.
60. Steinar Westin: UNEMPLOYMENT AND HEALTH: Medical and social consequences of a factory closure in a ten-year controlled follow-up study.
61. Ylva Sahlin: INJURY REGISTRATION, a tool for accident preventive work.
62. Helge Bjørnstad Pettersen: BIOSYNTHESIS OF COMPLEMENT BY HUMAN ALVEOLAR MACROPHAGES WITH SPECIAL REFERENCE TO SARCOIDOSIS.
63. Berit Schei: TRAPPED IN PAINFUL LOVE.
64. Lars J. Vatten: PROSPECTIVE STUDIES OF THE RISK OF BREAST CANCER IN A COHORT OF NORWEGIAN WOMAN.

1991

65. Kåre Bergh: APPLICATIONS OF ANTI-C5a SPECIFIC MONOCLONAL ANTIBODIES FOR THE ASSESSMENT OF COMPLEMENT ACTIVATION.
66. Svein Svenningsen: THE CLINICAL SIGNIFICANCE OF INCREASED FEMORAL ANTEVERSION.
67. Olbjørn Klepp: NONSEMINOMATOUS GERM CELL TESTIS CANCER: THERAPEUTIC OUTCOME AND PROGNOSTIC FACTORS.
68. Trond Sand: THE EFFECTS OF CLICK POLARITY ON BRAINSTEM AUDITORY EVOKED POTENTIALS AMPLITUDE, DISPERSION, AND LATENCY VARIABLES.
69. Kjetil B. Åsbakk: STUDIES OF A PROTEIN FROM PSORIATIC SCALE, PSO P27, WITH RESPECT TO ITS POTENTIAL ROLE IN IMMUNE REACTIONS IN PSORIASIS.
70. Arnulf Hestnes: STUDIES ON DOWN´S SYNDROME.
71. Randi Nygaard: LONG-TERM SURVIVAL IN CHILDHOOD LEUKEMIA.
72. Bjørn Hagen: THIO-TEPA.
73. Svein Anda: EVALUATION OF THE HIP JOINT BY COMPUTED TOMOGRAMPHY AND ULTRASONOGRAPHY.

1992

74. Martin Svartberg: AN INVESTIGATION OF PROCESS AND OUTCOME OF SHORT-TERM PSYCHODYNAMIC PSYCHOTHERAPY.
75. Stig Arild Slørdahl: AORTIC REGURGITATION.
76. Harold C Sexton: STUDIES RELATING TO THE TREATMENT OF SYMPTOMATIC NON-PSYCHOTIC PATIENTS.
77. Maurice B. Vincent: VASOACTIVE PEPTIDES IN THE OCULAR/FOREHEAD AREA.
78. Terje Johannessen: CONTROLLED TRIALS IN SINGLE SUBJECTS.
79. Turid Nilsen: PYROPHOSPHATE IN HEPATOCYTE IRON METABOLISM.
80. Olav Haraldseth: NMR SPECTROSCOPY OF CEREBRAL ISCHEMIA AND REPERFUSION IN RAT.
81. Eiliv Brenna: REGULATION OF FUNCTION AND GROWTH OF THE OXYNTIC MUCOSA.

1993

82. Gunnar Bovim: CERVICOGENIC HEADACHE.
83. Jarl Arne Kahn: ASSISTED PROCREATION.
84. Bjørn Naume: IMMUNOREGULATORY EFFECTS OF CYTOKINES ON NK CELLS.
85. Rune Wiseth: AORTIC VALVE REPLACEMENT.
86. Jie Ming Shen: BLOOD FLOW VELOCITY AND RESPIRATORY STUDIES.
87. Piotr Kruszewski: SUNCT SYNDROME WITH SPECIAL REFERENCE TO THE AUTONOMIC NERVOUS SYSTEM.
88. Mette Haase Moen: ENDOMETRIOSIS.
89. Anne Vik: VASCULAR GAS EMBOLISM DURING AIR INFUSION AND AFTER DECOMPRESSION IN PIGS.
90. Lars Jacob Stovner: THE CHIARI TYPE I MALFORMATION.
91. Kjell Å. Salvesen: ROUTINE ULTRASONOGRAPHY IN UTERO AND DEVELOPMENT IN CHILDHOOD.

1994

92. Nina-Beate Liabakk: DEVELOPMENT OF IMMUNOASSAYS FOR TNF AND ITS SOLUBLE RECEPTORS.
93. Sverre Helge Torp: *erb*B ONCOGENES IN HUMAN GLIOMAS AND MENINGIOMAS.
94. Olav M. Linaker: MENTAL RETARDATION AND PSYCHIATRY. Past and present.
95. Per Oscar Feet: INCREASED ANTIDEPRESSANT AND ANTIPANIC EFFECT IN COMBINED TREATMENT WITH DIXYRAZINE AND TRICYCLIC ANTIDEPRESSANTS.
96. Stein Olav Samstad: CROSS SECTIONAL FLOW VELOCITY PROFILES FROM TWO-DIMENSIONAL DOPPLER ULTRASOUND: Studies on early mitral blood flow.
97. Bjørn Backe: STUDIES IN ANTENATAL CARE.
98. Gerd Inger Ringdal: QUALITY OF LIFE IN CANCER PATIENTS.
99. Torvid Kiserud: THE DUCTUS VENOSUS IN THE HUMAN FETUS.
100. Hans E. Fjøsne: HORMONAL REGULATION OF PROSTATIC METABOLISM.
101. Eylert Brodtkorb: CLINICAL ASPECTS OF EPILEPSY IN THE MENTALLY RETARDED.
102. Roar Juul: PEPTIDERGIC MECHANISMS IN HUMAN SUBARACHNOID HEMORRHAGE.
103. Unni Syversen: CHROMOGRANIN A. Phsysiological and Clinical Role.

1995

104. Odd Gunnar Brakstad: THERMOSTABLE NUCLEASE AND THE *nuc* GENE IN THE DIAGNOSIS OF *Staphylococcus aureus* INFECTIONS.
105. Terje Engan: NUCLEAR MAGNETIC RESONANCE (NMR) SPECTROSCOPY OF PLASMA IN MALIGNANT DISEASE.
106. Kirsten Rasmussen: VIOLENCE IN THE MENTALLY DISORDERED.
107. Finn Egil Skjeldestad: INDUCED ABORTION: Timetrends and Determinants.
108. Roar Stenseth: THORACIC EPIDURAL ANALGESIA IN AORTOCORONARY BYPASS SURGERY.
109. Arild Faxvaag: STUDIES OF IMMUNE CELL FUNCTION *in mice infected with* MURINE RETROVIRUS.

1996

110. Svend Aakhus: NONINVASIVE COMPUTERIZED ASSESSMENT OF LEFT VENTRICULAR FUNCTION AND SYSTEMIC ARTERIAL PROPERTIES. Methodology and some clinical applications.
111. Klaus-Dieter Bolz: INTRAVASCULAR ULTRASONOGRAPHY.
112. Petter Aadahl: CARDIOVASCULAR EFFECTS OF THORACIC AORTIC CROSS-CLAMPING.
113. Sigurd Steinshamn: CYTOKINE MEDIATORS DURING GRANULOCYTOPENIC INFECTIONS.
114. Hans Stifoss-Hanssen: SEEKING MEANING OR HAPPINESS?
115. Anne Kvikstad: LIFE CHANGE EVENTS AND MARITAL STATUS IN RELATION TO RISK AND PROGNOSIS OF CANCER.
116. Torbjørn Grøntvedt: TREATMENT OF ACUTE AND CHRONIC ANTERIOR CRUCIATE LIGAMENT INJURIES. A clinical and biomechanical study.
117. Sigrid Hørven Wigers: CLINICAL STUDIES OF FIBROMYALGIA WITH FOCUS ON ETIOLOGY, TREATMENT AND OUTCOME.
118. Jan Schjøtt: MYOCARDIAL PROTECTION: Functional and Metabolic Characteristics of Two Endogenous Protective Principles.
119. Marit Martinussen: STUDIES OF INTESTINAL BLOOD FLOW AND ITS RELATION TO TRANSITIONAL CIRCULATORY ADAPATION IN NEWBORN INFANTS.
120. Tomm B. Müller: MAGNETIC RESONANCE IMAGING IN FOCAL CEREBRAL ISCHEMIA.
121. Rune Haaverstad: OEDEMA FORMATION OF THE LOWER EXTREMITIES.
122. Magne Børset: THE ROLE OF CYTOKINES IN MULTIPLE MYELOMA, WITH SPECIAL REFERENCE TO HEPATOCYTE GROWTH FACTOR.
123. Geir Smedslund: A THEORETICAL AND EMPIRICAL INVESTIGATION OF SMOKING, STRESS AND DISEASE: RESULTS FROM A POPULATION SURVEY.

1997

124. Torstein Vik: GROWTH, MORBIDITY, AND PSYCHOMOTOR DEVELOPMENT IN INFANTS WHO WERE GROWTH RETARDED *IN UTERO*.
125. Siri Forsmo: ASPECTS AND CONSEQUENCES OF OPPORTUNISTIC SCREENING FOR CERVICAL CANCER. Results based on data from three Norwegian counties.
126. Jon S. Skranes: CEREBRAL MRI AND NEURODEVELOPMENTAL OUTCOME IN VERY LOW BIRTH WEIGHT (VLBW) CHILDREN. A follow-up study of a geographically based year cohort of VLBW children at ages one and six years.
127. Knut Bjørnstad: COMPUTERIZED ECHOCARDIOGRAPHY FOR EVALUTION OF CORONARY ARTERY DISEASE.
128. Grethe Elisabeth Borchgrevink: DIAGNOSIS AND TREATMENT OF WHIPLASH/NECK SPRAIN INJURIES CAUSED BY CAR ACCIDENTS.
129. Tor Elsås: NEUROPEPTIDES AND NITRIC OXIDE SYNTHASE IN OCULAR AUTONOMIC AND SENSORY NERVES.
130. Rolf W. Gråwe: EPIDEMIOLOGICAL AND NEUROPSYCHOLOGICAL PERSPECTIVES ON SCHIZOPHRENIA.
131. Tonje Strømholm: CEREBRAL HAEMODYNAMICS DURING THORACIC AORTIC CROSSCLAMPING. An experimental study in pigs

1998

132. Martinus Bråten: STUDIES ON SOME PROBLEMS REALTED TO INTRAMEDULLARY NAILING OF FEMORAL FRACTURES.

133. Ståle Nordgård: PROLIFERATIVE ACTIVITY AND DNA CONTENT AS PROGNOSTIC INDICATORS IN ADENOID CYSTIC CARCINOMA OF THE HEAD AND NECK.
134. Egil Lien: SOLUBLE RECEPTORS FOR **TNF** AND **LPS**: RELEASE PATTERN AND POSSIBLE SIGNIFICANCE IN DISEASE.
135. Marit Bjørgaas: HYPOGLYCAEMIA IN CHILDREN WITH DIABETES MELLITUS
136. Frank Skorpen: GENETIC AND FUNCTIONAL ANALYSES OF DNA REPAIR IN HUMAN CELLS.
137. Juan A. Pareja: SUNCT SYNDROME. ON THE CLINICAL PICTURE. ITS DISTINCTION FROM OTHER, SIMILAR HEADACHES.
138. Anders Angelsen: NEUROENDOCRINE CELLS IN HUMAN PROSTATIC CARCINOMAS AND THE PROSTATIC COMPLEX OF RAT, GUINEA PIG, CAT AND DOG.
139. Fabio Antonaci: CHRONIC PAROXYSMAL HEMICRANIA AND HEMICRANIA CONTINUA: TWO DIFFERENT ENTITIES?
140. Sven M. Carlsen: ENDOCRINE AND METABOLIC EFFECTS OF METFORMIN WITH SPECIAL EMPHASIS ON CARDIOVASCULAR RISK FACTORES.

1999

141. Terje A. Murberg: DEPRESSIVE SYMPTOMS AND COPING AMONG PATIENTS WITH CONGESTIVE HEART FAILURE.
142. Harm-Gerd Karl Blaas: THE EMBRYONIC EXAMINATION. Ultrasound studies on the development of the human embryo.
143. Noèmi Becser Andersen:THE CEPHALIC SENSORY NERVES IN UNILATERAL HEADACHES. Anatomical background and neurophysiological evaluation.
144. Eli-Janne Fiskerstrand: LASER TREATMENT OF PORT WINE STAINS. A study of the efficacy and limitations of the pulsed dye laser. Clinical and morfological analyses aimed at improving the therapeutic outcome.
145. Bård Kulseng: A STUDY OF ALGINATE CAPSULE PROPERTIES AND CYTOKINES IN RELATION TO INSULIN DEPENDENT DIABETES MELLITUS.
146. Terje Haug: STRUCTURE AND REGULATION OF THE HUMAN UNG GENE ENCODING URACIL-DNA GLYCOSYLASE.
147. Heidi Brurok: MANGANESE AND THE HEART. A Magic Metal with Diagnostic and Therapeutic Possibilites.
148. Agnes Kathrine Lie: DIAGNOSIS AND PREVALENCE OF HUMAN PAPILLOMAVIRUS INFECTION IN CERVICAL INTRAEPITELIAL NEOPLASIA. Relationship to Cell Cycle Regulatory Proteins and HLA DQBI Genes.
149. Ronald Mårvik: PHARMACOLOGICAL, PHYSIOLOGICAL AND PATHOPHYSIOLOGICAL STUDIES ON ISOLATED STOMACS.
150. Ketil Jarl Holen: THE ROLE OF ULTRASONOGRAPHY IN THE DIAGNOSIS AND TREATMENT OF HIP DYSPLASIA IN NEWBORNS.
151. Irene Hetlevik: THE ROLE OF CLINICAL GUIDELINES IN CARDIOVASCULAR RISK INTERVENTION IN GENERAL PRACTICE.
152. Katarina Tunòn: ULTRASOUND AND PREDICTION OF GESTATIONAL AGE.
153. Johannes Soma: INTERACTION BETWEEN THE LEFT VENTRICLE AND THE SYSTEMIC ARTERIES.
154. Arild Aamodt: DEVELOPMENT AND PRE-CLINICAL EVALUATION OF A CUSTOM-MADE FEMORAL STEM.
155. Agnar Tegnander: DIAGNOSIS AND FOLLOW-UP OF CHILDREN WITH SUSPECTED OR KNOWN HIP DYSPLASIA.
156. Bent Indredavik: STROKE UNIT TREATMENT: SHORT AND LONG-TERM EFFECTS
157. Jolanta Vanagaite Vingen: PHOTOPHOBIA AND PHONOPHOBIA IN PRIMARY HEADACHES

2000

158. Ola Dalsegg Sæther: PATHOPHYSIOLOGY DURING PROXIMAL AORTIC CROSS-CLAMPING CLINICAL AND EXPERIMENTAL STUDIES
159. xxxxxxxxx (blind number)
160. Christina Vogt Isaksen: PRENATAL ULTRASOUND AND POSTMORTEM FINDINGS – A TEN YEAR CORRELATIVE STUDY OF FETUSES AND INFANTS WITH DEVELOPMENTAL ANOMALIES.
161. Holger Seidel: HIGH-DOSE METHOTREXATE THERAPY IN CHILDREN WITH ACUTE LYMPHOCYTIC LEUKEMIA: DOSE, CONCENTRATION, AND EFFECT CONSIDERATIONS.

162. Stein Hallan: IMPLEMENTATION OF MODERN MEDICAL DECISION ANALYSIS INTO CLINICAL DIAGNOSIS AND TREATMENT.
163. Malcolm Sue-Chu: INVASIVE AND NON-INVASIVE STUDIES IN CROSS-COUNTRY SKIERS WITH ASTHMA-LIKE SYMPTOMS.
164. Ole-Lars Brekke: EFFECTS OF ANTIOXIDANTS AND FATTY ACIDS ON TUMOR NECROSIS FACTOR-INDUCED CYTOTOXICITY.
165. Jan Lundbom: AORTOCORONARY BYPASS SURGERY: CLINICAL ASPECTS, COST CONSIDERATIONS AND WORKING ABILITY.
166. John-Anker Zwart: LUMBAR NERVE ROOT COMPRESSION, BIOCHEMICAL AND NEUROPHYSIOLOGICAL ASPECTS.
167. Geir Falck: HYPEROSMOLALITY AND THE HEART.
168. Eirik Skogvoll: CARDIAC ARREST Incidence, Intervention and Outcome.
169. Dalius Bansevicius: SHOULDER-NECK REGION IN CERTAIN HEADACHES AND CHRONIC PAIN SYNDROMES.
170. Bettina Kinge: REFRACTIVE ERRORS AND BIOMETRIC CHANGES AMONG UNIVERSITY STUDENTS IN NORWAY.
171. Gunnar Qvigstad: CONSEQUENCES OF HYPERGASTRINEMIA IN MAN
172. Hanne Ellekjær: EPIDEMIOLOGICAL STUDIES OF STROKE IN A NORWEGIAN POPULATION. INCIDENCE, RISK FACTORS AND PROGNOSIS
173. Hilde Grimstad: VIOLENCE AGAINST WOMEN AND PREGNANCY OUTCOME.
174. Astrid Hjelde: SURFACE TENSION AND COMPLEMENT ACTIVATION: Factors influencing bubble formation and bubble effects after decompression.
175. Kjell A. Kvistad: MR IN BREAST CANCER – A CLINICAL STUDY.
176. Ivar Rossvoll: ELECTIVE ORTHOPAEDIC SURGERY IN A DEFINED POPULATION. Studies on demand, waiting time for treatment and incapacity for work.
177. Carina Seidel: PROGNOSTIC VALUE AND BIOLOGICAL EFFECTS OF HEPATOCYTE GROWTH FACTOR AND SYNDECAN-1 IN MULTIPLE MYELOMA.

2001

178. Alexander Wahba: THE INFLUENCE OF CARDIOPULMONARY BYPASS ON PLATELET FUNCTION AND BLOOD COAGULATION – DETERMINANTS AND CLINICAL CONSEQUENSES
179. Marcus Schmitt-Egenolf: THE RELEVANCE OF THE MAJOR hISTOCOMPATIBILITY COMPLEX FOR THE GENETICS OF PSORIASIS
180. Odrun Arna Gederaas: BIOLOGICAL MECHANISMS INVOLVED IN 5-AMINOLEVULINIC ACID BASED PHOTODYNAMIC THERAPY
181. Pål Richard Romundstad: CANCER INCIDENCE AMONG NORWEGIAN ALUMINIUM WORKERS
182. Henrik Hjorth-Hansen: NOVEL CYTOKINES IN GROWTH CONTROL AND BONE DISEASE OF MULTIPLE MYELOMA
183. Gunnar Morken: SEASONAL VARIATION OF HUMAN MOOD AND BEHAVIOUR
184. Bjørn Olav Haugen: MEASUREMENT OF CARDIAC OUTPUT AND STUDIES OF VELOCITY PROFILES IN AORTIC AND MITRAL FLOW USING TWO- AND THREE-DIMENSIONAL COLOUR FLOW IMAGING
185. Geir Bråthen: THE CLASSIFICATION AND CLINICAL DIAGNOSIS OF ALCOHOL-RELATED SEIZURES
186. Knut Ivar Aasarød: RENAL INVOLVEMENT IN INFLAMMATORY RHEUMATIC DISEASE. A Study of Renal Disease in Wegener's Granulomatosis and in Primary Sjögren's Syndrome
187. Trude Helen Flo: RESEPTORS INVOLVED IN CELL ACTIVATION BY DEFINED URONIC ACID POLYMERS AND BACTERIAL COMPONENTS
188. Bodil Kavli: HUMAN URACIL-DNA GLYCOSYLASES FROM THE UNG GENE: STRUCTRUAL BASIS FOR SUBSTRATE SPECIFICITY AND REPAIR
189. Liv Thommesen: MOLECULAR MECHANISMS INVOLVED IN TNF- AND GASTRIN-MEDIATED GENE REGULATION
190. Turid Lingaas Holmen: SMOKING AND HEALTH IN ADOLESCENCE; THE NORD-TRØNDELAG HEALTH STUDY, 1995-97
191. Øyvind Hjertner: MULTIPLE MYELOMA: INTERACTIONS BETWEEN MALIGNANT PLASMA CELLS AND THE BONE MICROENVIRONMENT

192. Asbjørn Støylen: STRAIN RATE IMAGING OF THE LEFT VENTRICLE BY ULTRASOUND. FEASIBILITY, CLINICAL VALIDATION AND PHYSIOLOGICAL ASPECTS
193. Kristian Midthjell: DIABETES IN ADULTS IN NORD-TRØNDELAG. PUBLIC HEALTH ASPECTS OF DIABETES MELLITUS IN A LARGE, NON-SELECTED NORWEGIAN POPULATION.
194. Guanglin Cui: FUNCTIONAL ASPECTS OF THE ECL CELL IN RODENTS
195. Ulrik Wisløff: CARDIAC EFFECTS OF AEROBIC ENDURANCE TRAINING: HYPERTROPHY, CONTRACTILITY AND CALCUIM HANDLING IN NORMAL AND FAILING HEART
196. Øyvind Halaas: MECHANISMS OF IMMUNOMODULATION AND CELL-MEDIATED CYTOTOXICITY INDUCED BY BACTERIAL PRODUCTS
197. Tore Amundsen: PERFUSION MR IMAGING IN THE DIAGNOSIS OF PULMONARY EMBOLISM
198. Nanna Kurtze: THE SIGNIFICANCE OF ANXIETY AND DEPRESSION IN FATIQUE AND PATTERNS OF PAIN AMONG INDIVIDUALS DIAGNOSED WITH FIBROMYALGIA: RELATIONS WITH QUALITY OF LIFE, FUNCTIONAL DISABILITY, LIFESTYLE, EMPLOYMENT STATUS, CO-MORBIDITY AND GENDER
199. Tom Ivar Lund Nilsen: PROSPECTIVE STUDIES OF CANCER RISK IN NORD-TRØNDELAG: THE HUNT STUDY. Associations with anthropometric, socioeconomic, and lifestyle risk factors
200. Asta Kristine Håberg: A NEW APPROACH TO THE STUDY OF MIDDLE CEREBRAL ARTERY OCCLUSION IN THE RAT USING MAGNETIC RESONANCE TECHNIQUES

2002

201. Knut Jørgen Arntzen: PREGNANCY AND CYTOKINES
202. Henrik Døllner: INFLAMMATORY MEDIATORS IN PERINATAL INFECTIONS
203. Asta Bye: LOW FAT, LOW LACTOSE DIET USED AS PROPHYLACTIC TREATMENT OF ACUTE INTESTINAL REACTIONS DURING PELVIC RADIOTHERAPY. A PROSPECTIVE RANDOMISED STUDY.
204. Sylvester Moyo: STUDIES ON STREPTOCOCCUS AGALACTIAE (GROUP B STREPTOCOCCUS) SURFACE-ANCHORED MARKERS WITH EMPHASIS ON STRAINS AND HUMAN SERA FROM ZIMBABWE.
205. Knut Hagen: HEAD-HUNT: THE EPIDEMIOLOGY OF HEADACHE IN NORD-TRØNDELAG
206. Li Lixin: ON THE REGULATION AND ROLE OF UNCOUPLING PROTEIN-2 IN INSULIN PRODUCING ß-CELLS
207. Anne Hildur Henriksen: SYMPTOMS OF ALLERGY AND ASTHMA VERSUS MARKERS OF LOWER AIRWAY INFLAMMATION AMONG ADOLESCENTS
208. Egil Andreas Fors: NON-MALIGNANT PAIN IN RELATION TO PSYCHOLOGICAL AND ENVIRONTENTAL FACTORS. EXPERIENTAL AND CLINICAL STUDES OF PAIN WITH FOCUS ON FIBROMYALGIA
209. Pål Klepstad: MORPHINE FOR CANCER PAIN
210. Ingunn Bakke: MECHANISMS AND CONSEQUENCES OF PEROXISOME PROLIFERATOR-INDUCED HYPERFUNCTION OF THE RAT GASTRIN PRODUCING CELL
211. Ingrid Susann Gribbestad: MAGNETIC RESONANCE IMAGING AND SPECTROSCOPY OF BREAST CANCER
212. Rønnaug Astri Ødegård: PREECLAMPSIA – MATERNAL RISK FACTORS AND FETAL GROWTH
213. Johan Haux: STUDIES ON CYTOTOXICITY INDUCED BY HUMAN NATURAL KILLER CELLS AND DIGITOXIN
214. Turid Suzanne Berg-Nielsen: PARENTING PRACTICES AND MENTALLY DISORDERED ADOLESCENTS
215. Astrid Rydning: BLOOD FLOW AS A PROTECTIVE FACTOR FOR THE STOMACH MUCOSA. AN EXPERIMENTAL STUDY ON THE ROLE OF MAST CELLS AND SENSORY AFFERENT NEURONS

2003

216. Jan Pål Loennechen: HEART FAILURE AFTER MYOCARDIAL INFARCTION. Regional Differences, Myocyte Function, Gene Expression, and Response to Cariporide, Losartan, and Exercise Training.

217. Elisabeth Qvigstad: EFFECTS OF FATTY ACIDS AND OVER-STIMULATION ON INSULIN SECRETION IN MAN
218. Arne Åsberg: EPIDEMIOLOGICAL STUDIES IN HEREDITARY HEMOCHROMATOSIS: PREVALENCE, MORBIDITY AND BENEFIT OF SCREENING.
219. Johan Fredrik Skomsvoll: REPRODUCTIVE OUTCOME IN WOMEN WITH RHEUMATIC DISEASE. A population registry based study of the effects of inflammatory rheumatic disease and connective tissue disease on reproductive outcome in Norwegian women in 1967-1995.
220. Siv Mørkved: URINARY INCONTINENCE DURING PREGNANCY AND AFTER DELIVERY: EFFECT OF PELVIC FLOOR MUSCLE TRAINING IN PREVENTION AND TREATMENT
221. Marit S. Jordhøy: THE IMPACT OF COMPREHENSIVE PALLIATIVE CARE
222. Tom Christian Martinsen: HYPERGASTRINEMIA AND HYPOACIDITY IN RODENTS – CAUSES AND CONSEQUENCES
223. Solveig Tingulstad: CENTRALIZATION OF PRIMARY SURGERY FOR OVARAIN CANCER. FEASIBILITY AND IMPACT ON SURVIVAL
224. Haytham Eloqayli: METABOLIC CHANGES IN THE BRAIN CAUSED BY EPILEPTIC SEIZURES
225. Torunn Bruland: STUDIES OF EARLY RETROVIRUS-HOST INTERACTIONS – VIRAL DETERMINANTS FOR PATHOGENESIS AND THE INFLUENCE OF SEX ON THE SUSCEPTIBILITY TO FRIEND MURINE LEUKAEMIA VIRUS INFECTION
226. Torstein Hole: DOPPLER ECHOCARDIOGRAPHIC EVALUATION OF LEFT VENTRICULAR FUNCTION IN PATIENTS WITH ACUTE MYOCARDIAL INFARCTION
227. Vibeke Nossum: THE EFFECT OF VASCULAR BUBBLES ON ENDOTHELIAL FUNCTION
228. Sigurd Fasting: ROUTINE BASED RECORDING OF ADVERSE EVENTS DURING ANAESTHESIA – APPLICATION IN QUALITY IMPROVEMENT AND SAFETY
229. Solfrid Romundstad: EPIDEMIOLOGICAL STUDIES OF MICROALBUMINURIA. THE NORD-TRØNDELAG HEALTH STUDY 1995-97 (HUNT 2)
230. Geir Torheim: PROCESSING OF DYNAMIC DATA SETS IN MAGNETIC RESONANCE IMAGING
231. Catrine Ahlén: SKIN INFECTIONS IN OCCUPATIONAL SATURATION DIVERS IN THE NORTH SEA AND THE IMPACT OF THE ENVIRONMENT
232. Arnulf Langhammer: RESPIRATORY SYMPTOMS, LUNG FUNCTION AND BONE MINERAL DENSITY IN A COMPREHENSIVE POPULATION SURVEY. THE NORD-TRØNDELAG HEALTH STUDY 1995-97. THE BRONCHIAL OBSTRUCTION IN NORD-TRØNDELAG STUDY
233. Einar Kjelsås: EATING DISORDERS AND PHYSICAL ACTIVITY IN NON-CLINICAL SAMPLES
234. Arne Wibe: RECTAL CANCER TREATMENT IN NORWAY – STANDARDISATION OF SURGERY AND QUALITY ASSURANCE

2004

235. Eivind Witsø: BONE GRAFT AS AN ANTIBIOTIC CARRIER
236. Anne Mari Sund: DEVELOPMENT OF DEPRESSIVE SYMPTOMS IN EARLY ADOLESCENCE
237. Hallvard Lærum: EVALUATION OF ELECTRONIC MEDICAL RECORDS – A CLINICAL TASK PERSPECTIVE
238. Gustav Mikkelsen: ACCESSIBILITY OF INFORMATION IN ELECTRONIC PATIENT RECORDS; AN EVALUATION OF THE ROLE OF DATA QUALITY
239. Steinar Krokstad: SOCIOECONOMIC INEQUALITIES IN HEALTH AND DISABILITY. SOCIAL EPIDEMIOLOGY IN THE NORD-TRØNDELAG HEALTH STUDY (HUNT), NORWAY
240. Arne Kristian Myhre: NORMAL VARIATION IN ANOGENITAL ANATOMY AND MICROBIOLOGY IN NON-ABUSED PRESCHOOL CHILDREN
241. Ingunn Dybedal: NEGATIVE REGULATORS OF HEMATOPOIETEC STEM AND PROGENITOR CELLS
242. Beate Sitter: TISSUE CHARACTERIZATION BY HIGH RESOLUTION MAGIC ANGLE SPINNING MR SPECTROSCOPY
243. Per Arne Aas: MACROMOLECULAR MAINTENANCE IN HUMAN CELLS – REPAIR OF URACIL IN DNA AND METHYLATIONS IN DNA AND RNA

244. Anna Bofin: FINE NEEDLE ASPIRATION CYTOLOGY IN THE PRIMARY INVESTIGATION OF BREAST TUMOURS AND IN THE DETERMINATION OF TREATMENT STRATEGIES
245. Jim Aage Nøttestad: DEINSTITUTIONALIZATION AND MENTAL HEALTH CHANGES AMONG PEOPLE WITH MENTAL RETARDATION
246. Reidar Fossmark: GASTRIC CANCER IN JAPANESE COTTON RATS
247. Wibeke Nordhøy: MANGANESE AND THE HEART, INTRACELLULAR MR RELAXATION AND WATER EXCHANGE ACROSS THE CARDIAC CELL MEMBRANE

2005

248. Sturla Molden: QUANTITATIVE ANALYSES OF SINGLE UNITS RECORDED FROM THE HIPPOCAMPUS AND ENTORHINAL CORTEX OF BEHAVING RATS
249. Wenche Brenne Drøyvold: EPIDEMIOLOGICAL STUDIES ON WEIGHT CHANGE AND HEALTH IN A LARGE POPULATION. THE NORD-TRØNDELAG HEALTH STUDY (HUNT)
250. Ragnhild Støen: ENDOTHELIUM-DEPENDENT VASODILATION IN THE FEMORAL ARTERY OF DEVELOPING PIGLETS
251. Aslak Steinsbekk: HOMEOPATHY IN THE PREVENTION OF UPPER RESPIRATORY TRACT INFECTIONS IN CHILDREN
252. Hill-Aina Steffenach: MEMORY IN HIPPOCAMPAL AND CORTICO-HIPPOCAMPAL CIRCUITS
253. Eystein Stordal: ASPECTS OF THE EPIDEMIOLOGY OF DEPRESSIONS BASED ON SELF-RATING IN A LARGE GENERAL HEALTH STUDY (THE HUNT-2 STUDY)
254. Viggo Pettersen: FROM MUSCLES TO SINGING: THE ACTIVITY OF ACCESSORY BREATHING MUSCLES AND THORAX MOVEMENT IN CLASSICAL SINGING
255. Marianne Fyhn: SPATIAL MAPS IN THE HIPPOCAMPUS AND ENTORHINAL CORTEX
256. Robert Valderhaug: OBSESSIVE-COMPULSIVE DISORDER AMONG CHILDREN AND ADOLESCENTS: CHARACTERISTICS AND PSYCHOLOGICAL MANAGEMENT OF PATIENTS IN OUTPATIENT PSYCHIATRIC CLINICS
257. Erik Skaaheim Haug: INFRARENAL ABDOMINAL AORTIC ANEURYSMS – COMORBIDITY AND RESULTS FOLLOWING OPEN SURGERY
258. Daniel Kondziella: GLIAL-NEURONAL INTERACTIONS IN EXPERIMENTAL BRAIN DISORDERS
259. Vegard Heimly Brun: ROUTES TO SPATIAL MEMORY IN HIPPOCAMPAL PLACE CELLS
260. Kenneth McMillan: PHYSIOLOGICAL ASSESSMENT AND TRAINING OF ENDURANCE AND STRENGTH IN PROFESSIONAL YOUTH SOCCER PLAYERS
261. Marit Sæbø Indredavik: MENTAL HEALTH AND CEREBRAL MAGNETIC RESONANCE IMAGING IN ADOLESCENTS WITH LOW BIRTH WEIGHT
262. Ole Johan Kemi: ON THE CELLULAR BASIS OF AEROBIC FITNESS, INTENSITY-DEPENDENCE AND TIME-COURSE OF CARDIOMYOCYTE AND ENDOTHELIAL ADAPTATIONS TO EXERCISE TRAINING
263. Eszter Vanky: POLYCYSTIC OVARY SYNDROME – METFORMIN TREATMENT IN PREGNANCY
264. Hild Fjærtoft: EXTENDED STROKE UNIT SERVICE AND EARLY SUPPORTED DISCHARGE. SHORT AND LONG-TERM EFFECTS
265. Grete Dyb: POSTTRAUMATIC STRESS REACTIONS IN CHILDREN AND ADOLESCENTS
266. Vidar Fykse: SOMATOSTATIN AND THE STOMACH
267. Kirsti Berg: OXIDATIVE STRESS AND THE ISCHEMIC HEART: A STUDY IN PATIENTS UNDERGOING CORONARY REVASCULARIZATION
268. Björn Inge Gustafsson: THE SEROTONIN PRODUCING ENTEROCHROMAFFIN CELL, AND EFFECTS OF HYPERSEROTONINEMIA ON HEART AND BONE

2006

269. Torstein Baade Rø: EFFECTS OF BONE MORPHOGENETIC PROTEINS, HEPATOCYTE GROWTH FACTOR AND INTERLEUKIN-21 IN MULTIPLE MYELOMA
270. May-Britt Tessem: METABOLIC EFFECTS OF ULTRAVIOLET RADIATION ON THE ANTERIOR PART OF THE EYE
271. Anne-Sofie Helvik: COPING AND EVERYDAY LIFE IN A POPULATION OF ADULTS WITH HEARING IMPAIRMENT

272. Therese Standal: MULTIPLE MYELOMA: THE INTERPLAY BETWEEN MALIGNANT PLASMA CELLS AND THE BONE MARROW MICROENVIRONMENT
273. Ingvild Saltvedt: TREATMENT OF ACUTELY SICK, FRAIL ELDERLY PATIENTS IN A GERIATRIC EVALUATION AND MANAGEMENT UNIT – RESULTS FROM A PROSPECTIVE RANDOMISED TRIAL
274. Birger Henning Endreseth: STRATEGIES IN RECTAL CANCER TREATMENT – FOCUS ON EARLY RECTAL CANCER AND THE INFLUENCE OF AGE ON PROGNOSIS
275. Anne Mari Aukan Rokstad: ALGINATE CAPSULES AS BIOREACTORS FOR CELL THERAPY
276. Mansour Akbari: HUMAN BASE EXCISION REPAIR FOR PRESERVATION OF GENOMIC STABILITY
277. Stein Sundstrøm: IMPROVING TREATMENT IN PATIENTS WITH LUNG CANCER – RESULTS FROM TWO MULITCENTRE RANDOMISED STUDIES
278. Hilde Pleym: BLEEDING AFTER CORONARY ARTERY BYPASS SURGERY - STUDIES ON HEMOSTATIC MECHANISMS, PROPHYLACTIC DRUG TREATMENT AND EFFECTS OF AUTOTRANSFUSION
279. Line Merethe Oldervoll: PHYSICAL ACTIVITY AND EXERCISE INTERVENTIONS IN CANCER PATIENTS
280. Boye Welde: THE SIGNIFICANCE OF ENDURANCE TRAINING, RESISTANCE TRAINING AND MOTIVATIONAL STYLES IN ATHLETIC PERFORMANCE AMONG ELITE JUNIOR CROSS-COUNTRY SKIERS
281. Per Olav Vandvik: IRRITABLE BOWEL SYNDROME IN NORWAY, STUDIES OF PREVALENCE, DIAGNOSIS AND CHARACTERISTICS IN GENERAL PRACTICE AND IN THE POPULATION
282. Idar Kirkeby-Garstad: CLINICAL PHYSIOLOGY OF EARLY MOBILIZATION AFTER CARDIAC SURGERY
283. Linn Getz: SUSTAINABLE AND RESPONSIBLE PREVENTIVE MEDICINE. CONCEPTUALISING ETHICAL DILEMMAS ARISING FROM CLINICAL IMPLEMENTATION OF ADVANCING MEDICAL TECHNOLOGY
284. Eva Tegnander: DETECTION OF CONGENITAL HEART DEFECTS IN A NON-SELECTED POPULATION OF 42,381 FETUSES
285. Kristin Gabestad Nørsett: GENE EXPRESSION STUDIES IN GASTROINTESTINAL PATHOPHYSIOLOGY AND NEOPLASIA
286. Per Magnus Haram: GENETIC VS. AQUIRED FITNESS: METABOLIC, VASCULAR AND CARDIOMYOCYTE ADAPTATIONS
287. Agneta Johansson: GENERAL RISK FACTORS FOR GAMBLING PROBLEMS AND THE PREVALENCE OF PATHOLOGICAL GAMBLING IN NORWAY
288. Svein Artur Jensen: THE PREVALENCE OF SYMPTOMATIC ARTERIAL DISEASE OF THE LOWER LIMB
289. Charlotte Björk Ingul: QUANITIFICATION OF REGIONAL MYOCARDIAL FUNCTION BY STRAIN RATE AND STRAIN FOR EVALUATION OF CORONARY ARTERY DISEASE. AUTOMATED VERSUS MANUAL ANALYSIS DURING ACUTE MYOCARDIAL INFARCTION AND DOBUTAMINE STRESS ECHOCARDIOGRAPHY
290. Jakob Nakling: RESULTS AND CONSEQUENCES OF ROUTINE ULTRASOUND SCREENING IN PREGNANCY – A GEOGRAPHIC BASED POPULATION STUDY
291. Anne Engum: DEPRESSION AND ANXIETY – THEIR RELATIONS TO THYROID DYSFUNCTION AND DIABETES IN A LARGE EPIDEMIOLOGICAL STUDY
292. Ottar Bjerkeset: ANXIETY AND DEPRESSION IN THE GENERAL POPULATION: RISK FACTORS, INTERVENTION AND OUTCOME – THE NORD-TRØNDELAG HEALTH STUDY (HUNT)
293. Jon Olav Drogset: RESULTS AFTER SURGICAL TREATMENT OF ANTERIOR CRUCIATE LIGAMENT INJURIES – A CLINICAL STUDY
294. Lars Fosse: MECHANICAL BEHAVIOUR OF COMPACTED MORSELLISED BONE – AN EXPERIMENTAL IN VITRO STUDY
295. Gunilla Klensmeden Fosse: MENTAL HEALTH OF PSYCHIATRIC OUTPATIENTS BULLIED IN CHILDHOOD
296. Paul Jarle Mork: MUSCLE ACTIVITY IN WORK AND LEISURE AND ITS ASSOCIATION TO MUSCULOSKELETAL PAIN

297. Björn Stenström: LESSONS FROM RODENTS: I: MECHANISMS OF OBESITY SURGERY – ROLE OF STOMACH. II: CARCINOGENIC EFFECTS OF *HELICOBACTER PYLORI* AND SNUS IN THE STOMACH

2007

298. Haakon R. Skogseth: INVASIVE PROPERTIES OF CANCER – A TREATMENT TARGET ? IN VITRO STUDIES IN HUMAN PROSTATE CANCER CELL LINES
299. Janniche Hammer: GLUTAMATE METABOLISM AND CYCLING IN MESIAL TEMPORAL LOBE EPILEPSY
300. May Britt Drugli: YOUNG CHILDREN TREATED BECAUSE OF ODD/CD: CONDUCT PROBLEMS AND SOCIAL COMPETENCIES IN DAY-CARE AND SCHOOL SETTINGS
301. Arne Skjold: MAGNETIC RESONANCE KINETICS OF MANGANESE DIPYRIDOXYL DIPHOSPHATE (MnDPDP) IN HUMAN MYOCARDIUM. STUDIES IN HEALTHY VOLUNTEERS AND IN PATIENTS WITH RECENT MYOCARDIAL INFARCTION
302. Siri Malm: LEFT VENTRICULAR SYSTOLIC FUNCTION AND MYOCARDIAL PERFUSION ASSESSED BY CONTRAST ECHOCARDIOGRAPHY
303. Valentina Maria do Rosario Cabral Iversen: MENTAL HEALTH AND PSYCHOLOGICAL ADAPTATION OF CLINICAL AND NON-CLINICAL MIGRANT GROUPS
304. Lasse Løvstakken: SIGNAL PROCESSING IN DIAGNOSTIC ULTRASOUND: ALGORITHMS FOR REAL-TIME ESTIMATION AND VISUALIZATION OF BLOOD FLOW VELOCITY
305. Elisabeth Olstad: GLUTAMATE AND GABA: MAJOR PLAYERS IN NEURONAL METABOLISM
306. Lilian Leistad: THE ROLE OF CYTOKINES AND PHOSPHOLIPASE A_2S IN ARTICULAR CARTILAGE CHONDROCYTES IN RHEUMATOID ARTHRITIS AND OSTEOARTHRITIS
307. Arne Vaaler: EFFECTS OF PSYCHIATRIC INTENSIVE CARE UNIT IN AN ACUTE PSYCIATHRIC WARD
308. Mathias Toft: GENETIC STUDIES OF LRRK2 AND PINK1 IN PARKINSON'S DISEASE
309. Ingrid Løvold Mostad: IMPACT OF DIETARY FAT QUANTITY AND QUALITY IN TYPE 2 DIABETES WITH EMPHASIS ON MARINE N-3 FATTY ACIDS
310. Torill Eidhammer Sjøbakk: MR DETERMINED BRAIN METABOLIC PATTERN IN PATIENTS WITH BRAIN METASTASES AND ADOLESCENTS WITH LOW BIRTH WEIGHT
311. Vidar Beisvåg: PHYSIOLOGICAL GENOMICS OF HEART FAILURE: FROM TECHNOLOGY TO PHYSIOLOGY
312. Olav Magnus Søndenå Fredheim: HEALTH RELATED QUALITY OF LIFE ASSESSMENT AND ASPECTS OF THE CLINICAL PHARMACOLOGY OF METHADONE IN PATIENTS WITH CHRONIC NON-MALIGNANT PAIN
313. Anne Brantberg: FETAL AND PERINATAL IMPLICATIONS OF ANOMALIES IN THE GASTROINTESTINAL TRACT AND THE ABDOMINAL WALL
314. Erik Solligård: GUT LUMINAL MICRODIALYSIS
315. Elin Tollefsen: RESPIRATORY SYMPTOMS IN A COMPREHENSIVE POPULATION BASED STUDY AMONG ADOLESCENTS 13-19 YEARS. YOUNG-HUNT 1995-97 AND 2000-01; THE NORD-TRØNDELAG HEALTH STUDIES (HUNT)
316. Anne-Tove Brenne: GROWTH REGULATION OF MYELOMA CELLS
317. Heidi Knobel: FATIGUE IN CANCER TREATMENT – ASSESSMENT, COURSE AND ETIOLOGY
318. Torbjørn Dahl: CAROTID ARTERY STENOSIS. DIAGNOSTIC AND THERAPEUTIC ASPECTS
319. Inge-Andre Rasmussen jr.: FUNCTIONAL AND DIFFUSION TENSOR MAGNETIC RESONANCE IMAGING IN NEUROSURGICAL PATIENTS
320. Grete Helen Bratberg: PUBERTAL TIMING – ANTECEDENT TO RISK OR RESILIENCE ? EPIDEMIOLOGICAL STUDIES ON GROWTH, MATURATION AND HEALTH RISK BEHAVIOURS; THE YOUNG HUNT STUDY, NORD-TRØNDELAG, NORWAY
321. Sveinung Sørhaug: THE PULMONARY NEUROENDOCRINE SYSTEM. PHYSIOLOGICAL, PATHOLOGICAL AND TUMOURIGENIC ASPECTS
322. Olav Sande Eftedal: ULTRASONIC DETECTION OF DECOMPRESSION INDUCED VASCULAR MICROBUBBLES
323. Rune Bang Leistad: PAIN, AUTONOMIC ACTIVATION AND MUSCULAR ACTIVITY RELATED TO EXPERIMENTALLY-INDUCED COGNITIVE STRESS IN HEADACHE PATIENTS

324. Svein Brekke: TECHNIQUES FOR ENHANCEMENT OF TEMPORAL RESOLUTION IN THREE-DIMENSIONAL ECHOCARDIOGRAPHY
325. Kristian Bernhard Nilsen: AUTONOMIC ACTIVATION AND MUSCLE ACTIVITY IN RELATION TO MUSCULOSKELETAL PAIN
326. Anne Irene Hagen: HEREDITARY BREAST CANCER IN NORWAY. DETECTION AND PROGNOSIS OF BREAST CANCER IN FAMILIES WITH *BRCA1* GENE MUTATION
327. Ingebjørg S. Juel : INTESTINAL INJURY AND RECOVERY AFTER ISCHEMIA. AN EXPERIMENTAL STUDY ON RESTITUTION OF THE SURFACE EPITHELIUM, INTESTINAL PERMEABILITY, AND RELEASE OF BIOMARKERS FROM THE MUCOSA
328. Runa Heimstad: POST-TERM PREGNANCY
329. Jan Egil Afset: ROLE OF ENTEROPATHOGENIC *ESCHERICHIA COLI* IN CHILDHOOD DIARRHOEA IN NORWAY
330. Bent Håvard Hellum: *IN VITRO* INTERACTIONS BETWEEN MEDICINAL DRUGS AND HERBS ON CYTOCHROME P-450 METABOLISM AND P-GLYCOPROTEIN TRANSPORT
331. Morten André Høydal: CARDIAC DYSFUNCTION AND MAXIMAL OXYGEN UPTAKE MYOCARDIAL ADAPTATION TO ENDURANCE TRAINING

2008

332. Andreas Møllerløkken: REDUCTION OF VASCULAR BUBBLES: METHODS TO PREVENT THE ADVERSE EFFECTS OF DECOMPRESSION
333. Anne Hege Aamodt: COMORBIDITY OF HEADACHE AND MIGRAINE IN THE NORD-TRØNDELAG HEALTH STUDY 1995-97
334. Brage Høyem Amundsen: MYOCARDIAL FUNCTION QUANTIFIED BY SPECKLE TRACKING AND TISSUE DOPPLER ECHOCARDIOGRAPHY – VALIDATION AND APPLICATION IN EXERCISE TESTING AND TRAINING
335. Inger Anne Næss: INCIDENCE, MORTALITY AND RISK FACTORS OF FIRST VENOUS THROMBOSIS IN A GENERAL POPULATION. RESULTS FROM THE SECOND NORD-TRØNDELAG HEALTH STUDY (HUNT2)
336. Vegard Bugten: EFFECTS OF POSTOPERATIVE MEASURES AFTER FUNCTIONAL ENDOSCOPIC SINUS SURGERY
337. Morten Bruvold: MANGANESE AND WATER IN CARDIAC MAGNETIC RESONANCE IMAGING
338. Miroslav Fris: THE EFFECT OF SINGLE AND REPEATED ULTRAVIOLET RADIATION ON THE ANTERIOR SEGMENT OF THE RABBIT EYE
339. Svein Arne Aase: METHODS FOR IMPROVING QUALITY AND EFFICIENCY IN QUANTITATIVE ECHOCARDIOGRAPHY – ASPECTS OF USING HIGH FRAME RATE
340. Roger Almvik: ASSESSING THE RISK OF VIOLENCE: DEVELOPMENT AND VALIDATION OF THE BRØSET VIOLENCE CHECKLIST
341. Ottar Sundheim: STRUCTURE-FUNCTION ANALYSIS OF HUMAN ENZYMES INITIATING NUCLEOBASE REPAIR IN DNA AND RNA
342. Anne Mari Undheim: SHORT AND LONG-TERM OUTCOME OF EMOTIONAL AND BEHAVIOURAL PROBLEMS IN YOUNG ADOLESCENTS WITH AND WITHOUT READING DIFFICULTIES
343. Helge Garåsen: THE TRONDHEIM MODEL. IMPROVING THE PROFESSIONAL COMMUNICATION BETWEEN THE VARIOUS LEVELS OF HEALTH CARE SERVICES AND IMPLEMENTATION OF INTERMEDIATE CARE AT A COMMUNITY HOSPITAL COULD PROVIDE BETTER CARE FOR OLDER PATIENTS. SHORT AND LONG TERM EFFECTS
344. Olav A. Foss: "THE ROTATION RATIOS METHOD". A METHOD TO DESCRIBE ALTERED SPATIAL ORIENTATION IN SEQUENTIAL RADIOGRAPHS FROM ONE PELVIS
345. Bjørn Olav Åsvold: THYROID FUNCTION AND CARDIOVASCULAR HEALTH
346. Torun Margareta Melø: NEURONAL GLIAL INTERACTIONS IN EPILEPSY
347. Irina Poliakova Eide: FETAL GROWTH RESTRICTION AND PRE-ECLAMPSIA: SOME CHARACTERISTICS OF FETO-MATERNAL INTERACTIONS IN DECIDUA BASALIS
348. Torunn Askim: RECOVERY AFTER STROKE. ASSESSMENT AND TREATMENT; WITH FOCUS ON MOTOR FUNCTION
349. Ann Elisabeth Åsberg: NEUTROPHIL ACTIVATION IN A ROLLER PUMP MODEL OF CARDIOPULMONARY BYPASS. INFLUENCE ON BIOMATERIAL, PLATELETS AND COMPLEMENT

350. Lars Hagen: REGULATION OF DNA BASE EXCISION REPAIR BY PROTEIN INTERACTIONS AND POST TRANSLATIONAL MODIFICATIONS
351. Sigrun Beate Kjøtrød: POLYCYSTIC OVARY SYNDROME – METFORMIN TREATMENT IN ASSISTED REPRODUCTION
352. Steven Keita Nishiyama: PERSPECTIVES ON LIMB-VASCULAR HETEROGENEITY: IMPLICATIONS FOR HUMAN AGING, SEX, AND EXERCISE
353. Sven Peter Näsholm: ULTRASOUND BEAMS FOR ENHANCED IMAGE QUALITY
354. Jon Ståle Ritland: PRIMARY OPEN-ANGLE GLAUCOMA & EXFOLIATIVE GLAUCOMA. SURVIVAL, COMORBIDITY AND GENETICS
355. Sigrid Botne Sando: ALZHEIMER'S DISEASE IN CENTRAL NORWAY. GENETIC AND EDUCATIONAL ASPECTS
356. Parvinder Kaur: CELLULAR AND MOLECULAR MECHANISMS BEHIND METHYLMERCURY-INDUCED NEUROTOXICITY
357. Ismail Cüneyt Güzey: DOPAMINE AND SEROTONIN RECEPTOR AND TRANSPORTER GENE POLYMORPHISMS AND EXTRAPYRAMIDAL SYMPTOMS. STUDIES IN PARKINSON'S DISEASE AND IN PATIENTS TREATED WITH ANTIPSYCHOTIC OR ANTIDEPRESSANT DRUGS
358. Brit Dybdahl: EXTRA-CELLULAR INDUCIBLE HEAT-SHOCK PROTEIN 70 (Hsp70) – A ROLE IN THE INFLAMMATORY RESPONSE ?
359. Kristoffer Haugarvoll: IDENTIFYING GENETIC CAUSES OF PARKINSON'S DISEASE IN NORWAY
360. Nadra Nilsen: TOLL-LIKE RECEPTOR 2 –EXPRESSION, REGULATION AND SIGNALING
361. Johan Håkon Bjørngaard: PATIENT SATISFACTION WITH OUTPATIENT MENTAL HEALTH SERVICES – THE INFLUENCE OF ORGANIZATIONAL FACTORS.
362. Kjetil Høydal : EFFECTS OF HIGH INTENSITY AEROBIC TRAINING IN HEALTHY SUBJECTS AND CORONARY ARTERY DISEASE PATIENTS; THE IMPORTANCE OF INTENSITY,, DURATION AND FREQUENCY OF TRAINING.
363. Trine Karlsen: TRAINING IS MEDICINE: ENDURANCE AND STRENGTH TRAINING IN CORONARY ARTERY DISEASE AND HEALTH.
364. Marte Thuen: MANGANASE-ENHANCED AND DIFFUSION TENSOR MR IMAGING OF THE NORMAL, INJURED AND REGENERATING RAT VISUAL PATHWAY
365. Cathrine Broberg Vågbø: DIRECT REPAIR OF ALKYLATION DAMAGE IN DNA AND RNA BY 2-OXOGLUTARATE- AND IRON-DEPENDENT DIOXYGENASES
366. Arnt Erik Tjønna: AEROBIC EXERCISE AND CARDIOVASCULAR RISK FACTORS IN OVERWEIGHT AND OBESE ADOLESCENTS AND ADULTS
367. Marianne W. Furnes: FEEDING BEHAVIOR AND BODY WEIGHT DEVELOPMENT: LESSONS FROM RATS
368. Lene N. Johannessen: FUNGAL PRODUCTS AND INFLAMMATORY RESPONSES IN HUMAN MONOCYTES AND EPITHELIAL CELLS
369. Anja Bye: GENE EXPRESSION PROFILING OF *INHERITED* AND *ACQUIRED* MAXIMAL OXYGEN UPTAKE – RELATIONS TO THE METABOLIC SYNDROME.
370. Oluf Dimitri Røe: MALIGNANT MESOTHELIOMA: VIRUS, BIOMARKERS AND GENES. A TRANSLATIONAL APPROACH
371. Ane Cecilie Dale: DIABETES MELLITUS AND FATAL ISCHEMIC HEART DISEASE. ANALYSES FROM THE HUNT1 AND 2 STUDIES
372. Jacob Christian Hølen: PAIN ASSESSMENT IN PALLIATIVE CARE: VALIDATION OF METHODS FOR SELF-REPORT AND BEHAVIOURAL ASSESSMENT
373. Erming Tian: THE GENETIC IMPACTS IN THE ONCOGENESIS OF MULTIPLE MYELOMA
374. Ole Bosnes: KLINISK UTPRØVING AV NORSKE VERSJONER AV NOEN SENTRALE TESTER PÅ KOGNITIV FUNKSJON
375. Ola M. Rygh: 3D ULTRASOUND BASED NEURONAVIGATION IN NEUROSURGERY. A CLINICAL EVALUATION
376. Astrid Kamilla Stunes: ADIPOKINES, PEROXISOME PROFILERATOR ACTIVATED RECEPTOR (PPAR) AGONISTS AND SEROTONIN. COMMON REGULATORS OF BONE AND FAT METABOLISM
377. Silje Engdal: HERBAL REMEDIES USED BY NORWEGIAN CANCER PATIENTS AND THEIR ROLE IN HERB-DRUG INTERACTIONS
378. Kristin Offerdal: IMPROVED ULTRASOUND IMAGING OF THE FETUS AND ITS CONSEQUENCES FOR SEVERE AND LESS SEVERE ANOMALIES

379. Øivind Rognmo: HIGH-INTENSITY AEROBIC EXERCISE AND CARDIOVASCULAR HEALTH
380. Jo-Åsmund Lund: RADIOTHERAPY IN ANAL CARCINOMA AND PROSTATE CANCER

2009

381. Tore Grüner Bjåstad: HIGH FRAME RATE ULTRASOUND IMAGING USING PARALLEL BEAMFORMING
382. Erik Søndenaa: INTELLECTUAL DISABILITIES IN THE CRIMINAL JUSTICE SYSTEM
383. Berit Rostad: SOCIAL INEQUALITIES IN WOMEN'S HEALTH, HUNT 1984-86 AND 1995-97, THE NORD-TRØNDELAG HEALTH STUDY (HUNT)
384. Jonas Crosby: ULTRASOUND-BASED QUANTIFICATION OF MYOCARDIAL DEFORMATION AND ROTATION
385. Erling Tronvik: MIGRAINE, BLOOD PRESSURE AND THE RENIN-ANGIOTENSIN SYSTEM
386. Tom Christensen: BRINGING THE GP TO THE FOREFRONT OF EPR DEVELOPMENT
387. Håkon Bergseng: ASPECTS OF GROUP B STREPTOCOCCUS (GBS) DISEASE IN THE NEWBORN. EPIDEMIOLOGY, CHARACTERISATION OF INVASIVE STRAINS AND EVALUATION OF INTRAPARTUM SCREENING
388. Ronny Myhre: GENETIC STUDIES OF CANDIDATE TENE3S IN PARKINSON'S DISEASE
389. Torbjørn Moe Eggebø: ULTRASOUND AND LABOUR
390. Eivind Wang: TRAINING IS MEDICINE FOR PATIENTS WITH PERIPHERAL ARTERIAL DISEASE
391. Thea Kristin Våtsveen: GENETIC ABERRATIONS IN MYELOMA CELLS
392. Thomas Jozefiak: QUALITY OF LIFE AND MENTAL HEALTH IN CHILDREN AND ADOLESCENTS: CHILD AND PARENT PERSPECTIVES
393. Jens Erik Slagsvold: N-3 POLYUNSATURATED FATTY ACIDS IN HEALTH AND DISEASE – CLINICAL AND MOLECULAR ASPECTS
394. Kristine Misund: A STUDY OF THE TRANSCRIPTIONAL REPRESSOR ICER. REGULATORY NETWORKS IN GASTRIN-INDUCED GENE EXPRESSION
395. Franco M. Impellizzeri: HIGH-INTENSITY TRAINING IN FOOTBALL PLAYERS. EFFECTS ON PHYSICAL AND TECHNICAL PERFORMANCE
396. Kari Hanne Gjeilo: HEALTH-RELATED QUALITY OF LIFE AND CHRONIC PAIN IN PATIENTS UNDERGOING CARDIAC SURGERY
397. Øyvind Hauso: NEUROENDOCRINE ASPECTS OF PHYSIOLOGY AND DISEASE
398. Ingvild Bjellmo Johnsen: INTRACELLULAR SIGNALING MECHANISMS IN THE INNATE IMMUNE RESPONSE TO VIRAL INFECTIONS
399. Linda Tømmerdal Roten: GENETIC PREDISPOSITION FOR DEVELOPMENT OF PREEMCLAMPSIA – CANDIDATE GENE STUDIES IN THE HUNT (NORD-TRØNDELAG HEALTH STUDY) POPULATION
400. Trude Teoline Nausthaug Rakvåg: PHARMACOGENETICS OF MORPHINE IN CANCER PAIN
401. Hanne Lehn: MEMORY FUNCTIONS OF THE HUMAN MEDIAL TEMPORAL LOBE STUDIED WITH fMRI
402. Randi Utne Holt: ADHESION AND MIGRATION OF MYELOMA CELLS – IN VITRO STUDIES –
403. Trygve Solstad: NEURAL REPRESENTATIONS OF EUCLIDEAN SPACE
404. Unn-Merete Fagerli: MULTIPLE MYELOMA CELLS AND CYTOKINES FROM THE BONE MARROW ENVIRONMENT; ASPECTS OF GROWTH REGULATION AND MIGRATION
405. Sigrid Bjørnelv: EATING– AND WEIGHT PROBLEMS IN ADOLESCENTS, THE YOUNG HUNT-STUDY
406. Mari Hoff: CORTICAL HAND BONE LOSS IN RHEUMATOID ARTHRITIS. EVALUATING DIGITAL X-RAY RADIOGRAMMETRY AS OUTCOME MEASURE OF DISEASE ACTIVITY, RESPONSE VARIABLE TO TREATMENT AND PREDICTOR OF BONE DAMAGE
407. Siri Bjørgen: AEROBIC HIGH INTENSITY INTERVAL TRAINING IS AN EFFECTIVE TREATMENT FOR PATIENTS WITH CHRONIC OBSTRUCTIVE PULMONARY DISEASE
408. Susanne Lindqvist: VISION AND BRAIN IN ADOLESCENTS WITH LOW BIRTH WEIGHT
409. Torbjørn Hergum: 3D ULTRASOUND FOR QUANTITATIVE ECHOCARDIOGRAPHY

410. Jørgen Urnes: PATIENT EDUCATION IN GASTRO-OESOPHAGEAL REFLUX DISEASE. VALIDATION OF A DIGESTIVE SYMPTOMS AND IMPACT QUESTIONNAIRE AND A RANDOMISED CONTROLLED TRIAL OF PATIENT EDUCATION
411. Elvar Eyjolfsson: 13C NMRS OF ANIMAL MODELS OF SCHIZOPHRENIA
412. Marius Steiro Fimland: CHRONIC AND ACUTE NEURAL ADAPTATIONS TO STRENGTH TRAINING
413. Øyvind Støren: RUNNING AND CYCLING ECONOMY IN ATHLETES; DETERMINING FACTORS, TRAINING INTERVENTIONS AND TESTING
414. Håkon Hov: HEPATOCYTE GROWTH FACTOR AND ITS RECEPTOR C-MET. AUTOCRINE GROWTH AND SIGNALING IN MULTIPLE MYELOMA CELLS
415. Maria Radtke: ROLE OF AUTOIMMUNITY AND OVERSTIMULATION FOR BETA-CELL DEFICIENCY. EPIDEMIOLOGICAL AND THERAPEUTIC PERSPECTIVES
416. Liv Bente Romundstad: ASSISTED FERTILIZATION IN NORWAY: SAFETY OF THE REPRODUCTIVE TECHNOLOGY
417. Erik Magnus Berntsen: PREOPERATIV PLANNING AND FUNCTIONAL NEURONAVIGATION – WITH FUNCTIONAL MRI AND DIFFUSION TENSOR TRACTOGRAPHY IN PATIENTS WITH BRAIN LESIONS
418. Tonje Strømmen Steigedal: MOLECULAR MECHANISMS OF THE PROLIFERATIVE RESPONSE TO THE HORMONE GASTRIN
419. Vidar Rao: EXTRACORPOREAL PHOTOCHEMOTHERAPY IN PATIENTS WITH CUTANEOUS T CELL LYMPHOMA OR GRAFT-vs-HOST DISEASE
420. Torkild Visnes: DNA EXCISION REPAIR OF URACIL AND 5-FLUOROURACIL IN HUMAN CANCER CELL LINES

2010

421. John Munkhaugen: BLOOD PRESSURE, BODY WEIGHT, AND KIDNEY FUNCTION IN THE NEAR-NORMAL RANGE: NORMALITY, RISK FACTOR OR MORBIDITY ?
422. Ingrid Castberg: PHARMACOKINETICS, DRUG INTERACTIONS AND ADHERENCE TO TREATMENT WITH ANTIPSYCHOTICS: STUDIES IN A NATURALISTIC SETTING
423. Jian Xu: BLOOD-OXYGEN-LEVEL-DEPENDENT-FUNCTIONAL MAGNETIC RESONANCE IMAGING AND DIFFUSION TENSOR IMAGING IN TRAUMATIC BRAIN INJURY RESEARCH
424. Sigmund Simonsen: ACCEPTABLE RISK AND THE REQUIREMENT OF PROPORTIONALITY IN EUROPEAN BIOMEDICAL RESEARCH LAW. WHAT DOES THE REQUIREMENT THAT BIOMEDICAL RESEARCH SHALL NOT INVOLVE RISKS AND BURDENS DISPROPORTIONATE TO ITS POTENTIAL BENEFITS MEAN?
425. Astrid Woodhouse: MOTOR CONTROL IN WHIPLASH AND CHRONIC NON-TRAUMATIC NECK PAIN
426. Line Rørstad Jensen: EVALUATION OF TREATMENT EFFECTS IN CANCER BY MR IMAGING AND SPECTROSCOPY
427. Trine Moholdt: AEROBIC EXERCISE IN CORONARY HEART DISEASE
428. Øystein Olsen: ANALYSIS OF MANGANESE ENHANCED MRI OF THE NORMAL AND INJURED RAT CENTRAL NERVOUS SYSTEM
429. Bjørn H. Grønberg: PEMETREXED IN THE TREATMENT OF ADVANCED LUNG CANCER
430. Vigdis Schnell Husby: REHABILITATION OF PATIENTS UNDERGOING TOTAL HIP ARTHROPLASTY WITH FOCUS ON MUSCLE STRENGTH, WALKING AND AEROBIC ENDURANCE PERFORMANCE
431. Torbjørn Øien: CHALLENGES IN PRIMARY PREVENTION OF ALLERGY. THE PREVENTION OF ALLERGY AMONG CHILDREN IN TRONDHEIM (PACT) STUDY.
432. Kari Anne Indredavik Evensen: BORN TOO SOON OR TOO SMALL: MOTOR PROBLEMS IN ADOLESCENCE
433. Lars Adde: PREDICTION OF CEREBRAL PALSY IN YOUNG INFANTS. COMPUTER BASED ASSESSMENT OF GENERAL MOVEMENTS
434. Magnus Fasting: PRE- AND POSTNATAL RISK FACTORS FOR CHILDHOOD ADIPOSITY
435. Vivi Talstad Monsen: MECHANISMS OF ALKYLATION DAMAGE REPAIR BY HUMAN AlkB HOMOLOGUES
436. Toril Skandsen: MODERATE AND SEVERE TRAUMATIC BRAIN INJURY. MAGNETIC RESONANCE IMAGING FINDINGS, COGNITION AND RISK FACTORS FOR DISABILITY

437. Ingeborg Smidesang: ALLERGY RELATED DISORDERS AMONG 2-YEAR OLDS AND ADOLESCENTS IN MID-NORWAY – PREVALENCE, SEVERITY AND IMPACT. THE PACT STUDY 2005, THE YOUNG HUNT STUDY 1995-97
438. Vidar Halsteinli: MEASURING EFFICIENCY IN MENTAL HEALTH SERVICE DELIVERY: A STUDY OF OUTPATIENT UNITS IN NORWAY
439. Karen Lehrmann Ægidius: THE PREVALENCE OF HEADACHE AND MIGRAINE IN RELATION TO SEX HORMONE STATUS IN WOMEN. THE HUNT 2 STUDY
440. Madelene Ericsson: EXERCISE TRAINING IN GENETIC MODELS OF HEART FAILURE
441. Marianne Klokk: THE ASSOCIATION BETWEEN SELF-REPORTED ECZEMA AND COMMON MENTAL DISORDERS IN THE GENERAL POPULATION. THE HORDALAND HEALTH STUDY (HUSK)
442. Tomas Ottemo Stølen: IMPAIRED CALCIUM HANDLING IN ANIMAL AND HUMAN CARDIOMYOCYTES REDUCE CONTRACTILITY AND INCREASE ARRHYTHMIA POTENTIAL – EFFECTS OF AEROBIC EXERCISE TRAINING
443. Bjarne Hansen: ENHANCING TREATMENT OUTCOME IN COGNITIVE BEHAVIOURAL THERAPY FOR OBSESSIVE COMPULSIVE DISORDER: THE IMPORTANCE OF COGNITIVE FACTORS
444. Mona Løvlien: WHEN EVERY MINUTE COUNTS. FROM SYMPTOMS TO ADMISSION FOR ACUTE MYOCARDIAL INFARCTION WITH SPECIAL EMPHASIS ON GENDER DIFFERECES
445. Karin Margaretha Gilljam: DNA REPAIR PROTEIN COMPLEXES, FUNCTIONALITY AND SIGNIFICANCE FOR REPAIR EFFICIENCY AND CELL SURVIVAL
446. Anne Byriel Walls: NEURONAL GLIAL INTERACTIONS IN CEREBRAL ENERGY – AND AMINO ACID HOMEOSTASIS – IMPLICATIONS OF GLUTAMATE AND GABA
447. Cathrine Fallang Knetter: MECHANISMS OF TOLL-LIKE RECEPTOR 9 ACTIVATION
448. Marit Følsvik Svindseth: A STUDY OF HUMILIATION, NARCISSISM AND TREATMENT OUTCOME IN PATIENTS ADMITTED TO PSYCHIATRIC EMERGENCY UNITS
449. Karin Elvenes Bakkelund: GASTRIC NEUROENDOCRINE CELLS – ROLE IN GASTRIC NEOPLASIA IN MAN AND RODENTS
450. Kirsten Brun Kjelstrup: DORSOVENTRAL DIFFERENCES IN THE SPATIAL REPRESENTATION AREAS OF THE RAT BRAIN
451. Roar Johansen: MR EVALUATION OF BREAST CANCER PATIENTS WITH POOR PROGNOSIS
452. Rigmor Myran: POST TRAUMATIC NECK PAIN. EPIDEMIOLOGICAL, NEURORADIOLOGICAL AND CLINICAL ASPECTS
453. Krisztina Kunszt Johansen: GENEALOGICAL, CLINICAL AND BIOCHEMICAL STUDIES IN *LRRK2* – ASSOCIATED PARKINSON'S DISEASE
454. Pål Gjerden: THE USE OF ANTICHOLINERGIC ANTIPARKINSON AGENTS IN NORWAY. EPIDEMIOLOGY, TOXICOLOGY AND CLINICAL IMPLICATIONS
455. Else Marie Huuse: ASSESSMENT OF TUMOR MICROENVIRONMENT AND TREATMENT EFFECTS IN HUMAN BREAST CANCER XENOGRAFTS USING MR IMAGING AND SPECTROSCOPY
456. Khalid S. Ibrahim: INTRAOPERATIVE ULTRASOUND ASSESSMENT IN CORONARY ARTERY BYPASS SURGERY – WITH SPECIAL REFERENCE TO CORONARY ANASTOMOSES AND THE ASCENDING AORTA
457. Bjørn Øglænd: ANTHROPOMETRY, BLOOD PRESSURE AND REPRODUCTIVE DEVELOPMENT IN ADOLESCENCE OF OFFSPRING OF MOTHERS WHO HAD PREECLAMPSIA IN PREGNANCY
458. John Olav Roaldset: RISK ASSESSMENT OF VIOLENT, SUICIDAL AND SELF-INJURIOUS BEHAVIOUR IN ACUTE PSYCHIATRY – A BIO-PSYCHO-SOCIAL APPROACH
459. Håvard Dalen: ECHOCARDIOGRAPHIC INDICES OF CARDIAC FUNCTION – NORMAL VALUES AND ASSOCIATIONS WITH CARDIAC RISK FACTORS IN A POPULATION FREE FROM CARDIOVASCULAR DISEASE, HYPERTENSION AND DIABETES: THE HUNT 3 STUDY
460. Beate André: CHANGE CAN BE CHALLENGING. INTRODUCTION TO CHANGES AND IMPLEMENTATION OF COMPUTERIZED TECHNOLOGY IN HEALTH CARE
461. Latha Nrugham: ASSOCIATES AND PREDICTORS OF ATTEMPTED SUICIDE AMONG DEPRESSED ADOLESCENTS – A 6-YEAR PROSPECTIVE STUDY

462. Håvard Bersås Nordgaard: TRANSIT-TIME FLOWMETRY AND WALL SHEAR STRESS ANALYSIS OF CORONARY ARTERY BYPASS GRAFTS – A CLINICAL AND EXPERIMENTAL STUDY

Cotutelle with University of Ghent: Abigail Emily Swillens: A MULTIPHYSICS MODEL FOR IMPROVING THE ULTRASONIC ASSESSMENT OF LARGE ARTERIES

2011

463. Marte Helene Bjørk: DO BRAIN RHYTHMS CHANGE BEFORE THE MIGRAINE ATTACK? A LONGITUDINAL CONTROLLED EEG STUDY
464. Carl-Jørgen Arum: A STUDY OF UROTHELIAL CARCINOMA: GENE EXPRESSION PROFILING, TUMORIGENESIS AND THERAPIES IN ORTHOTOPIC ANIMAL MODELS
465. Ingunn Harstad: TUBERCULOSIS INFECTION AND DISEASE AMONG ASYLUM SEEKERS IN NORWAY. SCREENING AND FOLLOW-UP IN PUBLIC HEALTH CARE
466. Leif Åge Strand: EPIDEMIOLOGICAL STUDIES AMONG ROYAL NORWEGIAN NAVY SERVICEMEN. COHORT ESTABLISHMENT, CANCER INCIDENCE AND CAUSE-SPECIFIC MORTALITY
467. Katrine Høyer Holgersen: SURVIVORS IN THEIR THIRD DECADE AFTER THE NORTH SEA OIL RIG DISASTER OF 1980. LONG-TERM PERSPECTIVES ON MENTAL HEALTH
468. MarianneWallenius: PREGNANCY RELATED ASPECTS OF CHRONIC INFLAMMATORY ARTHRITIDES: DISEASE ONSET POSTPARTUM, PREGNANCY OUTCOMES AND FERTILITY. DATA FROM A NORWEGIAN PATIENT REGISTRY LINKED TO THE MEDICAL BIRTH REGISTRY OF NORWAY
469. Ole Vegard Solberg: 3D ULTRASOUND AND NAVIGATION – APPLICATIONS IN LAPAROSCOPIC SURGERY
470. Inga Ekeberg Schjerve: EXERCISE-INDUCED IMPROVEMENT OF MAXIMAL OXYGEN UPTAKE AND ENDOTHELIAL FUNCTION IN OBESE AND OVERWEIGHT INDIVIDUALS ARE DEPENDENT ON EXERCISE-INTENSITY
471. Eva Veslemøy Tyldum: CARDIOVASCULAR FUNCTION IN PREECLAMPSIA – WITH REFERENCE TO ENDOTHELIAL FUNCTION, LEFT VENTRICULAR FUNCTION AND PRE-PREGNANCY PHYSICAL ACTIVITY
472. Benjamin Garzón Jiménez de Cisneros: CLINICAL APPLICATIONS OF MULTIMODAL MAGNETIC RESONANCE IMAGING
473. Halvard Knut Nilsen: ASSESSING CODEINE TREATMENT TO PATIENTS WITH CHRONIC NON-MALIGNANT PAIN: NEUROPSYCHOLOGICAL FUNCTIONING, DRIVING ABILITY AND WEANING
474. Eiliv Brenner: GLUTAMATE RELATED METABOLISM IN ANIMAL MODELS OF SCHIZOPHRENIA
475. Egil Jonsbu: CHEST PAIN AND PALPITATIONS IN A CARDIAC SETTING; PSYCHOLOGICAL FACTORS, OUTCOME AND TREATMENT
476. Mona Høysæter Fenstad: GENETIC SUSCEPTIBILITY TO PREECLAMPSIA : STUDIES ON THE NORD-TRØNDELAG HEALTH STUDY (HUNT) COHORT, AN AUSTRALIAN/NEW ZEALAND FAMILY COHORT AND DECIDUA BASALIS TISSUE
477. Svein Erik Gaustad: CARDIOVASCULAR CHANGES IN DIVING: FROM HUMAN RESPONSE TO CELL FUNCTION
478. Karin Torvik: PAIN AND QUALITY OF LIFE IN PATIENTS LIVING IN NURSING HOMES
479. Arne Solberg: OUTCOME ASSESSMENTS IN NON-METASTATIC PROSTATE CANCER
480. Henrik Sahlin Pettersen: CYTOTOXICITY AND REPAIR OF URACIL AND 5-FLUOROURACIL IN DNA
481. Pui-Lam Wong: PHYSICAL AND PHYSIOLOGICAL CAPACITY OF SOCCER PLAYERS: EFFECTS OF STRENGTH AND CONDITIONING
482. Ole Solheim: ULTRASOUND GUIDED SURGERY IN PATIENTS WITH INTRACRANIAL TUMOURS
483. Sten Roar Snare: QUANTITATIVE CARDIAC ANALYSIS ALGORITHMS FOR POCKET-SIZED ULTRASOUND DEVICES
484. Marit Skyrud Bratlie: LARGE-SCALE ANALYSIS OF ORTHOLOGS AND PARALOGS IN VIRUSES AND PROKARYOTES
485. Anne Elisabeth F. Isern: BREAST RECONSTRUCTION AFTER MASTECTOMY – RISK OF RECURRENCE AFTER DELAYED LARGE FLAP RECONSTRUCTION – AESTHETIC OUTCOME, PATIENT SATISFACTION, QUALITY OF LIFE AND SURGICAL RESULTS;

HISTOPATHOLOGICAL FINDINGS AND FOLLOW-UP AFTER PROPHYLACTIC MASTECTOMY IN HEREDITARY BREAST CANCER

486. Guro L. Andersen: CEREBRAL PALSY IN NORWAY – SUBTYPES, SEVERITY AND RISK FACTORS
487. Frode Kolstad: CERVICAL DISC DISEASE – BIOMECHANICAL ASPECTS
488. Bente Nordtug: CARING BURDEN OF COHABITANTS LIVING WITH PARTNERS SUFFERING FROM CHRONIC OBSTRUCTIVE PULMONARY DISEASE OR DEMENTIA
489. Mariann Gjervik Heldahl: EVALUATION OF NEOADJUVANT CHEMOTHERAPY IN LOCALLY ADVANCED BREAST CANCER BASED ON MR METHODOLOGY
490. Lise Tevik Løvseth: THE SUBJECTIVE BURDEN OF CONFIDENTIALITY
491. Marie Hjelmseth Aune: INFLAMMATORY RESPONSES AGAINST GRAM NEGATIVE BACTERIA INDUCED BY TLR4 AND NLRP12
492. Tina Strømdal Wik: EXPERIMENTAL EVALUATION OF NEW CONCEPTS IN HIP ARTHROPLASTY
493. Solveig Sigurdardottir: CLINICAL ASPECTS OF CEREBRAL PALSY IN ICELAND. A POPULATION-BASED STUDY OF PRESCHOOL CHILDREN
494. Arne Reimers: CLINICAL PHARMACOKINETICS OF LAMOTRIGINE
495. Monica Wegling: KULTURMENNESKETS BYRDE OG SYKDOMMENS VELSIGNELSE. KAN MEDISINSK UTREDNING OG INTERVENSJON HA EN SELVSTENDIG FUNKSJON UAVHENGIG AV DET KURATIVE?
496. Silje Alvestad: ASTROCYTE-NEURON INTERACTIONS IN EXPERIMENTAL MESIAL TEMPORAL LOBE EPILEPSY – A STUDY OF UNDERLYING MECHANISMS AND POSSIBLE BIOMARKERS OF EPILEPTOGENESIS
497. Javaid Nauman: RESTING HEART RATE: A MATTER OF LIFE OR DEATH – PROSPECTIVE STUDIES OF RESTING HEART RATE AND CARDIOVASCULAR RISK (THE HUNT STUDY, NORWAY)
498. Thuy Nguyen: THE ROLE OF C-SRC TYROSINE KINASE IN ANTIVIRAL IMMUNE RESPONSES
499. Trine Naalsund Andreassen: PHARMACOKINETIC, PHARMACODYNAMIC AND PHARMACOGENETIC ASPECTS OF OXYCODONE TREATMENT IN CANCER PAIN
500. Eivor Alette Laugsand: SYMPTOMS IN PATIENTS RECEIVING OPIOIDS FOR CANCER PAIN – CLINICAL AND PHARMACOGENETIC ASPECTS
501. Dorthe Stensvold: PHYSICAL ACTIVITY, CARDIOVASCULAR HEALTH AND LONGEVITY IN PATIENTS WITH METABOLIC SYNDROME
502. Stian Thoresen Aspenes: PEAK OXYGEN UPTAKE AMONG HEALTHY ADULTS – CROSS-SECTIONAL DESCRIPTIONS AND PROSPECTIVE ANALYSES OF PEAK OXYGEN UPTAKE, PHYSICAL ACTIVITY AND CARDIOVASCULAR RISK FACTORS IN HEALTHY ADULTS (20-90 YEARS)
503. Reidar Alexander Vigen: PATHOBIOLOGY OF GASTRIC CARCINOIDS AND ADENOCARCINOMAS IN RODENT MODELS AND PATIENTS. STUDIES OF GASTROCYSTOPLASTY, GENDER-RELATED FACTORS, AND AUTOPHAGY
504. Halvard Høilund-Kaupang: MODELS AND METHODS FOR INVESTIGATION OF REVERBERATIONS IN NONLINEAR ULTRASOUND IMAGING
505. Audhild Løhre: WELLBEING AMONG SCHOOL CHILDREN IN GRADES 1-10: PROMOTING AND ADVERSE FACTORS
506. Torgrim Tandstad: VOX POPULI. POPULATION-BASED OUTCOME STUDIES IN TESTICULAR CANCER
507. Anna Brenne Grønskag: THE EPIDEMIOLOGY OF HIP FRACTURES AMONG ELDERLY WOMEN IN NORD-TRØNDELAG. HUNT 1995-97, THE NORD-TRØNDELAG HEALTH STUDY
508. Kari Ravndal Risnes: BIRTH SIZE AND ADULT MORTALITY: A SYSTEMATIC REVIEW AND A LONG-TERM FOLLOW-UP OF NEARLY 40 000 INDIVIDUALS BORN AT ST. OLAV UNIVERSITY HOSPITAL IN TRONDHEIM 1920-1960
509. Hans Jakob Bøe: LONG-TERM POSTTRAUMATIC STRESS AFTER DISASTER – A CONTROLLED STUDY OF SURVIVORS' HEALTH 27 YEARS AFTER THE CAPSIZED NORTH SEA OIL RIG
510. Cathrin Barbara Canto, Cotutelle with University of Amsterdam: LAYER SPECIFIC INTEGRATIVE PROPERTIES OF ENTORHINAL PRINCIPAL NEURONS
511. Ioanna Sandvig: THE ROLE OF OLFACTORY ENSHEATHING CELLS, MRI, AND BIOMATERIALS IN TRANSPLANT-MEDIATED CNS REPAIR

512. Karin Fahl Wader: HEPATOCYTE GROWTH FACTOR, C-MET AND SYNDECAN-1 IN MULTIPLE MYELOMA
513. Gerd Tranø: FAMILIAL COLORECTAL CANCER
514. Bjarte Bergstrøm: INNATE ANTIVIRAL IMMUNITY – MECHANISMS OF THE RIG-I-MEDIATED RESPONSE
515. Marie Søfteland Sandvei: INCIDENCE, MORTALITY, AND RISK FACTORS FOR ANEURYSMAL SUBARACHNOID HEMORRHAGE. PROSPECTIVE ANALYZES OF THE HUNT AND TROMSØ STUDIES
516. Mary-Elizabeth Bradley Eilertsen: CHILDREN AND ADOLESCENTS SURVIVING CANCER: PSYCHOSOCIAL HEALTH, QUALITY OF LIFE AND SOCIAL SUPPORT
517. Takaya Saito: COMPUTATIONAL ANALYSIS OF REGULATORY MECHANISM AND INTERACTIONS OF MICRORNAS

Godkjent for disputas, publisert post mortem: Eivind Jullumstrø: COLORECTAL CANCER AT LEVANGER HOSPITAL 1980-2004

518. Christian Gutvik: A PHYSIOLOGICAL APPROACH TO A NEW DECOMPRESSION ALGORITHM USING NONLINEAR MODEL PREDICTIVE CONTROL
519. Ola Storrø: MODIFICATION OF ADJUVANT RISK FACTOR BEHAVIOURS FOR ALLERGIC DISEASE AND ASSOCIATION BETWEEN EARLY GUT MICROBIOTA AND ATOPIC SENSITIZATION AND ECZEMA. EARLY LIFE EVENTS DEFINING THE FUTURE HEALTH OF OUR CHILDREN
520. Guro Fanneløb Giskeødegård: IDENTIFICATION AND CHARACTERIZATION OF PROGNOSTIC FACTORS IN BREAST CANCER USING MR METABOLOMICS
521. Gro Christine Christensen Løhaugen: BORN PRETERM WITH VERY LOW BIRTH WEIGHT – NEVER ENDING COGNITIVE CONSEQUENCES?
522. Sigrid Nakrem: MEASURING QUALITY OF CARE IN NURSING HOMES – WHAT MATTERS?
523. Brita Pukstad: CHARACTERIZATION OF INNATE INFLAMMATORY RESPONSES IN ACUTE AND CHRONIC WOUNDS

2012

524. Hans Wasmuth: ILEAL POUCHES
525. Inger Økland: BIASES IN SECOND-TRIMESTER ULTRASOUND DATING RELATED TO PREDICTION MODELS AND FETAL MEASUREMENTS