

One Norwegian national tool for estimating date of delivery and fetal age

The Norwegian Directorate of Health recommends the use of one national term prediction model, eSnurra, to determine the date of delivery and fetal age. This model has the best precision and the smallest systematic error, and is best in taking natural variation into account. The use of one single method nationwide helps avoid geographic variations, provides women seeking abortion with equality before the law, gives identical assessments of pre-term births and ensures an unambiguous assessment in post-term pregnancies.

Two Norwegian models for ultrasound-based determination of the estimated date of delivery (EDD) are currently available: eSnurra and Terminhjulet. The choice of method to use is not inconsequential. The Norwegian Directorate of Health recommends eSnurra. Terminhjulet suffers from a systematic bias that may cause too many pregnancies to be considered post-term, and some women may be given an erroneous EDD and an inaccurate fetal age. The outcome of an application for abortion may depend on which model is used. This could also entail significant consequences in terms of the treatment offered to women with threatening extreme pre-term delivery.

eSnurra is a population-based model developed at the National Centre for Fetal Medicine at St. Olavs Hospital, Trondheim, Norway, the Department of Medical Informatics at the University of Oslo, Norway, and the Norwegian Institute of Public Health (1). Stavanger University Hospital, Norway, has been involved in the evaluations (2).

Fetal age and estimated date of delivery – the problem of bias

Terminhjulet was introduced in 2004 on the basis of studies of a selected study population consisting of 650 women from Haukeland University Hospital, Bergen, Norway (3, 4). The authors claimed that Terminhjulet would provide better predictions of EDD than the old Snurra model. It has later transpired that this claim was incorrect. Terminhjulet was developed using the same statistical methods as the old Snurra, and suffers from similar flaws. Had its developers evaluated Terminhjulet with the same thoroughness shown in their evaluation of the old Snurra model (5), the systematic bias inherent in Terminhjulet would have been discovered.

eSnurra is a population-based term prediction model, and was developed from measurements in approximately 40 000 pregnant women. A population-based method will invariably be more robust than traditional sample-based models. eSnurra is adapted to the population of pregnant women as they are actually encountered in clinical practice, not as they appear in a selected study population. Research documentation shows that

eSnurra produces more correct results than traditional methods such as Terminhjulet with regard to both the EDD and the fetal age (2).

In contrast to population-based models, the traditional models *must* be validated prospectively before they can be applied to pregnant women. It was during such a

«Research documentation shows that eSnurra produces more correct results than traditional term prediction models such as Terminhjulet»

validation of the old Snurra method in 1997 that the problem of a systematic bias was scientifically described (6).

Fetal age and EDD always need to be considered in conjunction. In Terminhjulet, the EDD is derived from the age, in eSnurra the age is derived from the date of delivery. Evaluations (7–9) show that Terminhjulet's EDDs are unstable, with a varying bias over the weeks it was developed to cover – pregnancy weeks 13–23. This gives rise to significant variations, for example in the proportion of post-term deliveries, depending on the week in which the ultrasound examination is performed.

The problem of the bias in Terminhjulet originates in its *estimation of age*. The variable to be estimated – the gestational age counted from the last menstruation – is also the variable used to choose the date of the ultrasound examination in the study population. Moreover, the distribution of these examinations in the study population differs significantly from the distribution where the method is applied, i.e. in pregnant women at 18 weeks of gestation. The selection of data points based on the variable to be predicted (age) gives rise to a systematic bias in the *estimation of age*.

This leads to a corresponding bias in the *prediction of the EDD* and is thereby the cause of the documented systematic error in the Terminhjulet model. This mechanism has been described, especially by Økland and collaborators in 2012 (7), and it is an unfortunate consequence of the statistical method used in the traditional term prediction models, such as the old Snurra tool and Terminhjulet. A direct confirmation of this can be seen in Figure 2 of the article by Økland and collaborators (7).

Documentation

Ebbing and collaborators (10) have criticised the Norwegian Directorate of Health for its recommendation to apply eSnurra nationwide. The Directorate's recommendations are based on comprehensive scientific documentation, assessments by the Norwegian Knowledge Centre for the Health Services and expertise in academic communities in Norway and abroad, including with regard to the statistical methodology.

The Norwegian Knowledge Centre for the Health Services pointed out that one possible validation strategy to assess different term prediction models consists in comparing the characteristics of the methods in question through measurements within one and the same population. This was done in a PhD thesis in 2012 (2), and the work has been presented in three international publications (7–9). The PhD thesis undertook validation studies of three different term prediction models: Snurra (phased out in 2007), Terminhjulet and eSnurra. The methods were validated, and the consequences of their differences were shown with data from a total of 73 409 pregnancies from three different Norwegian regions.

The old Snurra model and Terminhjulet both showed a varying systematic bias in the prediction of EDD over the range of inclusion – pregnancy weeks 14–23. The bias amounted to a maximum of five days, equivalent in all three populations. When using eSnurra, the EDD corresponded to the actual median date of delivery, also equivalent in all three populations and without any systematic bias. The methods were validated on their precision in predicting fetal age, and eSnurra scored highest (7).

The commentary by Ebbing and collabo-

rators in the Journal of the Norwegian Medical Association (10) contains a number of fallacies: The developers of eSnurra are criticised for having used the population from which the method was developed in the evaluation of how well the EDD is predicted, and for using the same population to evaluate Terminhjulet. This is a fallacy. It is correct to evaluate the results in the same population from which the model was developed, as long as it is population-based.

It is also pointed out that a systematic review article gave a positive assessment of the quality of the studies of age determination and growth on which Terminhjulet is based (11). However, the presentation by Ebbing and collaborators is incorrect. The systematic review article is *not* a study of pregnancy dating models, but of fetal growth curves for use in later pregnancy.

According to Ebbing and collaborators, it must be expected that the EDD that was predicted with eSnurra would influence the health services to converge on this prediction. This is a fallacy. The pregnancies were dated from the old Snurra method, and the women were delivered long before the evaluation of eSnurra was initiated. They also refer to a critical comment on the statistical population-based approach that underlies eSnurra (12), but fail to mention the refutation of this criticism (13). This is quite remarkable.

It is wrong to claim that the due date is of no clinical importance. The fact that few births occur exactly on the EDD is not crucial. What is important is that the model has no systematic bias that causes the ultrasound-based EDD of the entire population to be shifted, as it is with Terminhjulet.

A thoroughly validated method

Since 1986, Norway has provided well-organised services for routine ultrasound examinations of all pregnant women in pregnancy weeks 17–19. Nearly all accept the offer of this service. The examination, including an estimation of the date of delivery and fetal age, is primarily undertaken by midwives who have all undergone the same specialist training. This practice has been documented as being of good quality in more than 70 000 pregnant women from three locations in Norway (7–9).

Ebbing and collaborators refer to recommendations from the Norwegian Gynaecological Association's manual, which contradict the well-validated services currently provided in Norway. This manual recommends that a first-trimester dating examination should not be amended by later ultrasound examinations. The manual does not

specify the model to be used for fetal age determination or the competence required for the ultrasound operators.

If we were to change practice from second- to first-trimester dating examinations, the examinations would not be accomplished in the same systematic manner as they are today. This involves a risk of reduced quality. Despite the emphasis placed on an evidence-based approach, no mention is made of how a new procedure should be evaluated. If the organisation of ultrasound examinations and pregnancy dating is to be changed in Norway, the model to be used first needs to be validated. No such validation has been made. A recent Danish study, however, has compared first- and second-trimester dating. Its conclusion is that validation must be undertaken beforehand to avoid erroneous predictions (14).

Professional coordination

The Norwegian Directorate of Health is charged with collating knowledge and experience, setting national norms and ensuring provision of equal and adequate health services to the population. Norwegian health legislation stipulates that the Directorate of Health is the only agency mandated to develop national medical guidelines and manuals. This is done in collaboration with the medical communities, which also fill an important role by developing evidence-based recommendations, supplementing the national guidelines from the Norwegian Directorate of Health. These are all key contributions to the totality of norms for good medical practice.

In the efforts undertaken to clarify which model can best determine the EDD and fetal age, the Norwegian Gynaecological Association has repeatedly been invited to meetings and been informed about the basis for the conclusions drawn by the Directorate of Health.

The Norwegian Ministry of Health and Care Services has requested the Norwegian Directorate of Health to enact the measures necessary to implement its recommendation. The medical directors of the regional health authorities have been informed and have sent a recommendation to use eSnurra to their healthcare providers.

The Directorate of Health expects Norwegian doctors and midwives to comply with the recommendation by collaborating well on the use of the most precise tool. This helps provide the best possible health services to pregnant women.

Torunn Janbu
torunn.janbu@helsedir.no

Torunn Janbu (born 1954), MD, PhD, specialist in general and orthopaedic surgery, head of the Department of Hospital Services, Norwegian Directorate of Health.

The author has completed the ICMJE form and declares no conflicts of interest.

References

- 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.
- Økland I. Biases in second-trimester ultrasound dating related to prediction models and foetal measurements. Doktoravhandling. Trondheim: Norges teknisk-naturvitenskapelige universitet, 2012: 4.
- Johnsen SL, Rasmussen S, Sollien R et al. Foetal age assessment based on ultrasound head biometry and the effect of maternal and foetal factors. *Acta Obstet Gynecol Scand* 2004; 83: 716–23.
- Johnsen SL, Rasmussen S, Sollien R et al. Foetal 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: 725–33.
- Kiserud T, Rasmussen S. Terminbestemmelse ved hjelp av ultralyd – kan metoden bli bedre? *Tidsskr Nor Lægeforen* 1999; 119: 4331–4.
- Tunón K, Eik-Nes SH, Grøttum P. The impact of foetal, maternal and external factors on prediction of the day of delivery by the use of ultrasound. *Ultrasound Obstet Gynecol* 1998; 11: 99–103.
- Økland I, Nakling J, Gjessing HK et al. Advantages of the population-based approach to pregnancy dating: results from 23,020 ultrasound examinations. *Ultrasound Obstet Gynecol* 2012; 39: 563–8.
- Økland I, Gjessing HK, Grøttum P et al. A new population-based term prediction model vs. two traditional sample-based models: validation on 9046 ultrasound examinations. *Ultrasound Obstet Gynecol* 2011; 37: 207–13.
- Økland I, Gjessing HK, Grøttum P et al. Biases of traditional term prediction models: results from different sample-based models evaluated on 41,343 ultrasound examinations. *Ultrasound Obstet Gynecol* 2010; 36: 728–34.
- Ebbing C, Johnsen SL, Kessler J et al. Helsedirektoratet gir feil anbefaling om bestemmelse av fosteralder. *Tidsskr Nor Lægeforen* 2015; 135: 740–1.
- Ioannou C, Talbot K, Ohuma E et al. Systematic review of methodology used in ultrasound studies aimed at creating charts of foetal size. *BJOG* 2012; 119: 1425–39.
- Kiserud T, Johnsen SL, Rasmussen S. Re: A direct method for ultrasound prediction of day of delivery: a new, population-based approach. Problems of accounting for a retrospective selection. *Ultrasound Obstet Gynecol* 2008; 31: 225.
- Kiserud T, Johnsen SL, Rasmussen S. Re: A direct method for ultrasound prediction of day of delivery: a new, population-based approach. Problems of accounting for a retrospective selection. *Ultrasound Obstet Gynecol* 2008; 31: 225–7.
- Näslund Thagaard I, Krebs L, Lausten-Thomsen U et al. Dating of pregnancy in first versus second trimester in relation to post-term birth rate: a cohort study. *PLoS One* 2016; 11: e0147109.

Received 14 November 2015, first revision submitted 27 February 2016, accepted 18 April 2016. Editor: Martine Rostadmo.