The Journal of Maternal-Fetal and Neonatal Medicine, August 2010; 23(8): 944–947

LETTER TO THE EDITOR

Comment and reply on: Prediction of the date of delivery based on first trimester ultrasound measurements: an independent method from estimated date of conception

We read with interest the paper 'Prediction of the date of delivery based on first trimester ultrasound measurements: an independent method from estimated date of conception' by Salomon et al. [1]. It is very interesting to see the median-based direct prediction method that we developed for second trimester measurements [2,3] being applied to first trimester data. Although the size of their study is much smaller than ours, it illustrates the strength of our population-based method; when high-quality registry data are available, the direct method provides a prediction model adapted to the clinical setting that the data were collected under, whether it is first- or second trimester routine scans.

To avoid obscuring the central concept, it should be stressed that the essence of the authors' approach is precisely the one developed in our paper, namely a direct estimation of median (and other centiles) of remaining time of pregnancy. This is the basis for deriving all the clinically useful estimates such as term prediction and prediction interval. The obvious statistical approach to the estimation is then to use a non-linear quantile regression model. A plethora of such models is available in the literature, and both the local linear quantile regression (LLQR) [4] employed by us and the spline-smoothed quantile regression employed by the authors fall in this category. Our reason for avoiding the spline smoothing approach was that it is inherently polynomial. This is clearly seen, for instance, in Figure 1 [1], where the spline approach ends up with a single knot at the median crown-rump length (CRL) value, and straight lines on each side, which may seem a bit odd. The LLQR method is to a larger extent truly non-linear in its kernel smoothing approach. However, we believe that such technical differences are of secondary importance to the overarching principle of direct median prediction.

As the authors point out, centiles of remaining time is a potentially useful tool in pregnancy management, and one of the advantages of the direct prediction method. In Figures 1, 2, 5, and 6 [2], we provided values for the 10th, 25th, 50th, 75th, and 90th centiles, avoiding the most extreme centiles for which the precision is more questionable. The authors are correct in claiming that the LLQR method does not provide means (meaning, presumably, a formula) of computing centiles. We fail to see, however, that the authors' suggested spline method provides such means. Regardless, a specific formula is not really needed. For any practical application of the model, the required centiles can be estimated from raw data and tabulated. Furthermore, since the shape of the residual distribution in our second trimester predictions is quite stable over the prediction range, useful and stable centile estimates can easily be obtained from the overall distribution (Figure 3), even for the more extreme centiles.

The authors refer to our study [2] as using 'less accurate second trimester measurements'. The conventional thinking is that in the first trimester the variability in fetal size is smaller, thus estimates made during the first trimester should be more precise than second trimester estimates. However, since the primary outcome is time of birth, not time of conception, this is not as obvious as it may seem. In fact, this is precisely one of the important empirical questions that can be addressed with our direct prediction method. If direct prediction curves were constructed both for first- and second trimester measurements, preferably on the same population, their precisions could be compared. This would effectively avoid uncertainties arising from comparing different populations and different dating formulas. Indeed, in Figures 1, 2, 5, and 6 [2] there are no obvious signs of increased precision at the lower end of the prediction range, which goes down to about week 13. Unfortunately, the authors miss the opportunity to study this in more detail and only refer to two previous publications. Their first reference is Taipale and Hiilesmaa [5], which in fact does not contain data after week 16; the paper only points onward to a paper on gestational age by Bergsjø et al. [6], which has no data on ultrasound at all, using only the last menstrual period (LMP). Their second reference is Saltvedt et al. [7], which is

ISSN 1476-7058 print/ISSN 1476-4954 online © 2010 Informa UK Ltd.
DOI: 10.3109/14767050903420309

more relevant in that it contains both early and late dating scans. However, it is based on only 363 *in vitro* fertilized pregnancies, which may or may not be representative of a full population. It evaluates a selection of dating formulas developed using the traditional indirect approach (based on the estimated day of conception), and not all the formulas are applied in the same trimester as they are developed for, which opens for further confusion. In fact, Saltvedt et al. emphasize in their discussion that '... we need a formula that performs well throughout our dating interval ...., and the technique must be the same as that used when the formula was designed.' This is precisely what the direct prediction method can provide. It may well be the case that first trimester scans provide higher precision than second trimester scans, but it would be very useful to see this verified scientifically, which could be done with our direct prediction method.

The authors do not present the frequency distributions of their ultrasound measurements, only the total number of measurements. The actual distributions would provide important information. While estimation of the median curve is likely to be stable, to estimate other centiles, in particular the extreme 1st and 99th centiles, a large amount of data are required. Clearly, out of the 3738 examinations in their study, <40 would be expected to fall below the 1st centile curve. If the frequency distribution of the CRL values peaks around, say, CRL = 65 mm, it is likely that very little, if any, data are available to estimate the 1st centile at the outer ends of the CRL range. To deal with this, the authors must apply a high degree of smoothing to the centile values and extrapolate to the outer ends. Furthermore, since not even the raw (unsmoothed) centiles are shown, it is very hard to assess the precision of the extreme centile estimates reported in Table IV and Figure 1 [1]. In fact, with more than 10 times their study size we noticed (Figures 5 and 6 [2]) that even the 10th and 90th raw centiles were not very stable at the ends of the range of inclusion.

Hand-in-hand with the problem of term prediction is the question of establishing a 'correct' gestational age. If term is predicted by the direct method, while gestational age still is being established using the traditional indirect approach, this would lead to two different time scales which are not necessarily mutually consistent. For instance, for a specific child the gestational age and the predicted remaining time might sum to different values, say 282 and 283, depending on at what age the child was given the ultrasound examination. It seems to be simpler, more robust and more consistent simply to define gestational age as median total length of pregnancy minus the predicted number of remaining days. In our prediction system we have decided to use 283 days as the total length, since this is the median 'certain LMP'-based pregnancy length in our data. Using this strategy, it is only the number 283 that needs to be computed from LMP. Once an ultrasound measurement has been made, the term prediction tables can be used to also compute the gestational age at the time of measurement.

In conclusion, we welcome the authors' application of our direct prediction approach to first trimester measurements. We believe our approach opens for the development of population-based prediction curves over a large range of measurements and populations. While we fail to see any particular advantages of the alternative quantile smoothing approach suggested by the authors, the advantages of our direct prediction method are still obvious: it obtains fully ultrasound-based estimates of median birth term, and allows the precision of these estimates to be assessed and quantified directly.

Håkon K. Gjessing[1], Per Grøttum[2], Inger Økland[3], & Sturla H. Eik-Nes[4]

[1]*Division of Epidemiology, Norwegian Institute of Public Health, Oslo, Norway,*
[2]*Department of Informatics, University of Oslo, Oslo, Norway,*
[3]*Department of Obstetrics and Gynecology, Stavanger University Hospital, Stavanger,Norway, and*
[4]*National Center for Fetal Medicine, St. Olavs University Hospital, Trondheim, Norway*

Correspondence: Håkon K. Gjessing, Norwegian Institute of Public Health, PO box 4404 Nydalen, N-0403 Oslo, Norway. E-mail: hakon.gjessing@fhi.no

References

1. Salomon LJ, Pizzi C, Gasparrini A, Bernard J-P, 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):1–9.
2. 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.
3. Eik-Nes SH, Blaas HG, Grøttum P, Gjessing HK. Predicting remaining time of pregnancy – a new approach to the prediction of day of delivery. Fifteenth World Congress on Ultrasound in Obstetrics and Gynecology, Vancouver, Canada; 2005.
4. Yu K, Jones MC. Local linear quantile regression. J Am Stat Assoc 1998;93:228–237.
5. Taipale P, Hiilesmaa V. Predicting delivery date by ultrasound and last menstrual period in early gestation. Obstet Gynecol 2001;97:189–194.

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

Author's reply

We are thankful to Gjessing et al. for their interesting comments on our paper 'Prediction of the date of delivery based on first trimester ultrasound measurements: an independent method from estimated date of conception', recently published on-line in this journal [1].

Gjessing et al. raised interesting issues regarding the methodological approaches we used to develop our predictions, focussing on the comparison with the methods they applied to a different and larger sample of foetal measurements [2], extensively cited in our paper. In particular, they commented the differences between our and their choices on statistical models (based on fully parametric and semi-parametric smoothing) and the prediction ability of measurements taken at first or second trimester, respectively.

Before addressing the specific comments, we are pleased to explain why we disagree with Gjessing et al. when they argue that they 'fail to see any particular advantage of the alternative quantile smoothing approach' we developed. These advantages are called comparability and reproducibility. Given that no method could be considered perfect to answer a scientific question, the availability of different analytic strategies offers a way to compare the benefits and weaknesses of each approach, strengthening the plausibility of the findings if the results are consistent between different choices. This issue is closely related to the second advantage: our spline-based approach [3] is performed through functions currently implemented in the major statistical packages, for example Stata (StataCorp, LP, TX) with *rc_pline* or *mkspline*, R (R Development Core Team) with *ns*, and SAS (SAS Institute) with *proc transreg*. The availability of these functions gives the opportunity to replicate our method in other populations, and to compare it against potential alternative solutions. To our knowledge, the local linear quantile regression (LLQR) [4] proposed by Gjessing et al. is not directly implemented in any major statistical software, and they needed to develop their own routine which is not available to other researchers.

Several different smoothing techniques have been proposed, ranging from non-parametric, semi-parametric and fully parametric methods. In spite of these differences, all these approaches share the common goal to describe a non-linear relationship between the predictor and the outcome, with the choice of the correct degree of smoothness guided by the trade-off between flexibility and stability of the function. In this sense, the spline-based method is not very different from the LLQR: in the former, the degree of flexibility is chosen by the number and location of the knots, while in the latter by the weight function based on the normal distribution, with its standard deviation expressing the bandwidth of the kernel. We chose to place the knots at equally spaced quantiles (the default in all the packages), and control the number of knots by cross-validation. Gjessing et al. reported to set the bandwidth to 1 in the central part of the predictor distribution, and ranging from 2 in the highest to 3 in the lowest values. However, they fail to report how they chose these values, and therefore how they controlled the flexibility of the curve. Our selection method preferred the model with only one knot, indicating that the degree of non-linearity is actually low, as showed by the Figures 1–4 in our paper [1]. These findings are indeed very similar to the pattern estimated by Gjessing and collaborator in Figures 1 and 2 [2].

Gjessing et al. did not include the estimated curve beyond 10th and 90th percentiles, arguing that any estimate would be based on few observations, especially at the boundaries of the predictor distribution. One of the benefits of every smoothing technique is that the model is able to 'borrow information' from the adjacent observations, in a way inversely proportional to the degree of flexibility allowed to the curve. In practice, the estimate is more stable (but more prone to bias) including less knots or increasing the bandwidth. This allows estimating the curve at the boundaries, where few observations are available. In addition, we use natural cubic splines [3], functions which are forced to assume a linear shape beyond the boundaries, and therefore more stable at the tails of the predictor distribution, where less observations are present. Anyway, we agree with Gjessing et al. that our estimates for the more extreme percentiles are less certain than that for the median. As we stated in the Discussion, the standard error for the model parameters is not included as a prediction error in order to define the distribution of the date of delivery, relying on the fact that, at least for the less extreme percentiles, this is negligible if compared to the natural variability of the phenomenon. Anyway, this is not true for the more extreme percentiles, and the incorporation of this additional source of uncertainty might represent an issue for future research.

In conclusion, we hope that the method we proposed will be replicated in other populations, and possibly improved to address the limitations described by our colleagues. Anyway, we deem that our well-implemented and reproducible method represents a useful tool for gestational age assessment and date of delivery prediction.

Laurent J. Salomon[1], C. Pizzi[2], A. Gasparrini[3], J.-P. Bernard[1], & Y. Ville[1]

[1]*Maternité, Hôpital Necker-Enfants Malades, Université Paris Descartes, AP-HP, 75015 Paris, France,* [2]*Medical Statistics Unit, Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, Keppel Street, London, WC1E 7HT, UK, and* [3]*Public and Environmental Health Research Unit, Department of Public Health and Policy, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK*

Correspondence: Dr Laurent J. Salomon, MD, PhD, *CHU Necker-Enfants Malades. APHP* 149, rue de Sèvres 75015 Paris. Tel: +33-1-44-49-40-30. E-mail: laurentsalomon@gmail.com

References

1. Salomon LJ, et al. 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 2009;July 14:1–9 [Epub ahead of print].
2. Gjessing HK, Grottum P, Eik-Nes S-H. A direct method for ultrasound prediction of day of delivery: a new, population-based approach. Ultrasound Obstet Gynecol 2007;30(1):19–27.
3. Hastie T, Tibishirani R. Generalized additive models. New York, NY: Chapman and Hall Inc.; 1990.
4. Yu K, Lu Z. Quantile regression: application and current research areas. The Statistician 2003;53:331–350.