Journal of Gynecology Obstetrics and Human Reproduction 48 (2019) 141–142

Available online at
ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com/en

Editorial

Machine learning to anticipate delivery room activity?

The restructuring of obstetric care provision in France, with the closure of many maternity units and a concentration of activity [1], combined with recent labor management guidelines [2] likely to extend the length of labor, has led to an increased pressure on delivery rooms and raise organizational challenge. The important variability in activity with phases of over activity, well known to obstetrician gynecologists who take shifts, are problematic in terms of the adequacy of needs and means and generate medical risk and users insatisfaction.

The regulation of activity through the extensive use of labor induction is proposed by some, especially since a 39-week labor induction strategy for low-risk nulliparous women who would agree is now supported by the results of a recently published trial [3]. However, the time between the beginning of induction and delivery is highly variable and is even more unpredictable. In addition, the time spent in the delivery room could be increased overall [4]. The organizational problems posed by the unequal distribution of births over time would therefore not be solved by this type of strategy.

A model powered by relevant data from pregnant women from the active line of a maternity unit, that would predict the number of women that could give birth in the next few days and allow for anticipation, would therefore be more relevant.

As of today, an expected due date is estimated for each woman according to the same calculus method. The day of pregnancy beginning is estimated with the measurement of the craniocaudal (crown-rump) length of the first trimester ultrasound. Then 287 days (41 weeks) are added to this date to compute the expected due date. This expected due date is almost never the actual date of delivery. Mongelli et al. [5] show that only 4% of women with a singleton pregnancy would deliver on their expected due date whereas 70% would give birth in a 10 days wide interval. This gap between the expected due date and the actual date of delivery is too wide to use the expected due date in an organizational perspective: it is not precise enough to anticipate activity in delivery rooms. Updating the expected due date for each woman all along the pregnancy duration by taking into account events and observations which have an influence on pregnancies duration could be of great help to reduce estimation variability.

The objectives for a finer-grained estimation of expected due date are first to anticipate the activity peaks to lessen the probability of delay in care because of under-sized material and human resources and second, to identify phases of low activity to optimize human resources.

Facing the limitations of current expected due date calculus we have sought to develop a 'big data oriented' predictive model to estimate the actual delivery date and predict the activity in the delivery ward.

This model was developed from data of the maternity unit Notre Dame de Bon Secours of the Paris Saint Joseph Hospital. We aimed at developing models to estimate the remaining time of a pregnancy. The final goal was to introduce a methodology framework to anticipate the number of pregnant women under hospital supervision who will give birth in the next 24 h using data gathered during their follow-up appointments at the maternity unit. Women who have given birth in 2017 at the maternity constituted our training set (n = 3370) and those who have given birth between January and June 2018 was our testing set (n = 1730).

Recent advances in machine learning and in calculus power allows us to handle high dimensional datasets. The main standpoint in the recent' Big Data for Healthcare' trend is to take all the data we have on patients and to process it without prejudice on what could be medically relevant and what is not. From this point of view conferring model conception to neither medical nor paramedical experts but to a' naive' data scientist is supposed to avoid any pre-existing influence on feature importance, which would lead to a biased analysis. However, because of the existence of non-obvious outliers that rely on expert knowledge to classify them and the need of this expertise to structure patient information, data scientists and physicians are working on this project hand to hand.

Considering the nature of prenatal follow-up, we developed three models to predict pregnancy length. The first gives an estimation of the due date at the very beginning of the pregnancy as it uses only the pre-pregnancy socio-demographic data (age, parity, weight before pregnancy, type of health coverage, country of birth, type of habitation . . .) collected during their first prenatal visit at the hospital. The second powered by medical data collected during the first prenatal visit after 32 weeks gestational age (weight gain, blood pressure . . .) updates the estimation. The last model additionally integrates to the data used in model 1 and 2, the data collected at all subsequent visits and gives an estimation of the day of delivery.

The best model found for predicting the number of remaining pregnancy days is a Support Vector Regression with a non-linear kernel (Radial Basis Function). Prediction by 41 weeks of amenorrhea has a mean absolute error (MAE) of 9 days on actual date of delivery predicting, while our best model MAE is just under 6 days. These

http://dx.doi.org/10.1016/j.jogoh.2018.11.012

model performances are encouraging as they are superior to those of the old estimation formula. However, they remain not good enough for an operational exploitation: absolute error on the predicted number of birth per day going from 3 to 2.3 births. To be usable in a facility with an annual activity of 3500 delivery, this model should be able to provide an estimate of the number of deliveries with an absolute error that would not exceed 1.

Given the nature of prenatal care and current patient electronic files, paths for pregnancy duration model improvement are narrow. First, structuration of patient electronic files triggers numerous data losses. Lots of features of interest gathered during follow-ups consist on open text fields. Given its variability, information stored as open text is hard to exploit automatically and statistical treatment becomes really tough. It is easily understandable why some information has to be reported as open text: when there is a massive number of possible answers or when further development is needed, aiming at asking practitioners to structure their answers is illusory and unpractical. Detecting keywords and semantics is possible but demands time-consuming expert knowledge. Research on automatic understanding of semantics is currently undergoing and is progressing but for now on information should be gathered on open text only if no other format is viable. For instance, data on cervix examination, fundal height or contraction presence are typical of what could be gathered in structured fields. From this perspective a massive work is to be done on hospitals' information systems to make their patient's data usable.

Second, in some areas, most of the prenatal follow-up is actually done outside hospitals. In France, according to the Perinatal National Study in 2016 the main practitioner responsible for the first six-month monitoring is an ob-gyn outside of hospital environment for more than 50% of pregnant women [6]. It means that for a vast majority of women, a large part of the data won't be available from the hospital database. After 32 weeks of gestation, 87% of pregnant women come at the hospital at least twice before the day they give birth and only 40% come at least three times. Thus, temporal evolution of pregnancy metrics could only be observed twice during the pregnancy: not enough to deploy state of the art time series models.

This methodological approach is brand new as it relies on the development of big data analysis and electronic patient files which has recently become widespread. Hospitals gather every day huge amounts of data on their patients, which constitute a real wealth for statisticians, data analysts and end users of the models. The point is now to make this raw data exploitable. It is very likely that by introducing a little more structuring in the medical files, semantic analysis approaches and time series analyses, this project could step up to the next level in terms of performance and eventually significantly improve resources allocations and risk management in delivery rooms.

References

[1] Pilkington H, Blondel B, Carayol M, Breart G, Zeitlin J. Impact of maternity unit closures on access to obstetrical care: the French experience between 1998 and 2003. Soc Sci Med 2008;67(November (10)):1521–9.

[2] Dupont C, Carayol M, Le Ray C, Barasinski C, Beranger R, Burguet A, et al. Oxytocin administration during spontaneous labour: guidelines for clinical practice. Guidelines short text. Gynecol Obstet Fertil Senol 2017;45(January (1)):56–61.

[3] Grobman WA, Rice MM, Reddy UM, Tita ATN, Silver RM, Mallett G, et al. Labor induction versus expectant management in low-risk nulliparous women. N Engl J Med 2018;379(August (6)):513–23.

[4] Greene MF. Choices in managing full-term pregnancy. N Engl J Med 2018;379 (August (6)):580–1.

[5] Mongelli M, Wilcox M, Gardosi J. Estimating the date of confinement: ultrasonographic biometry versus certain menstrual dates. Am J Obstet Gynecol 1996;174(January (1 Pt 1)):278–81.

[6] Blondel B, Gonzalez L, Coulm B, Bonnet C, Raynaud P, Vanhaesebrouck A, et al. Enquête nationale périnatale: rapport 2016. 2017.

Albane de l'Espinay[a,b]

[a] *Medical Information Department, Paris Saint Joseph Hospital, Paris, France*

[b] *Ecole polytechnique, Paris Saclay University, Palaiseau, France*

Catherine Sauvage[a,b]

[a] *Medical Information Department, Paris Saint Joseph Hospital, Paris, France*

[b] *Maternity Unit, Paris Saint Joseph Hospital, DHU Risk and Pregnancy, Paris, France*

Anne Buronfosse
Olivier Billuart

Medical Information Department, Paris Saint Joseph Hospital, Paris, France

Jean-Patrick Lajonchère

Executive Management, Paris Saint Joseph Hospital, Paris, France

Elie Azria[a,b,*]

[a] *Maternity Unit, Paris Saint Joseph Hospital, DHU Risk and Pregnancy, Paris, France*

[b] *UMR1153, Obstetrical, Perinatal and Pediatric Epidemiology Research Team (EPOPé), Paris Descartes University, Paris, France*

* Corresponding author at: Maternité du Groupe Hospitalier Paris Saint Joseph, 185 rue Raymond Losserand, 75674 Paris Cedex 14, France.

E-mail address: eazria@hpsj.fr (E. Azria).

Received 28 November 2018

Available online 1 December 2018