Article Type: Main Research Article

Cervical dilatation over time is a poor predictor of severe adverse birth outcomes: a diagnostic accuracy study

Joao Paulo Souza[1]*, Olufemi T Oladapo[1], Bukola Fawole[2], Kidza Mugerwa[3], Rodrigo Reis[4], Francisco Barbosa-Junior[4], Livia Oliveira-Ciabati[4], Domingos Alves[4], A Metin Gülmezoglu[1].

[1] UNDP/UNFPA/UNICEF/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction (HRP), Department of Reproductive Health and Research, World Health Organization, Geneva, Switzerland

[2] Department of Obstetrics and Gynaecology, College of Medicine, University of Ibadan, Ibadan, Nigeria

[3] Department of Obstetrics and Gynaecology, Makerere University, Kampala, Uganda

[4] Center for Information and Informatics for Health (CIIS), Department of Social Medicine, Ribeirão Preto Medical School, University of São Paulo, Ribeirão Preto, Brazil

* Correspondence: jp.souza@usp.br

Running title

The diagnostic accuracy of labour curves and the partograph.

Abstract

Objective: To assess the accuracy of the WHO partograph alert line and other candidate predictors in the identification of women at risk of developing severe adverse birth outcomes. Design: a facility-based, multicentre, prospective cohort study. Setting: 13 maternity hospitals located in Nigeria and Uganda. Population: 9,995 women with spontaneous onset of labour presenting at cervical dilatation

of ≤6 cm or undergoing induction of labour. **Methods:** Research assistants collected data on socio-demographic, anthropometric, obstetric and medical characteristics of study participants at hospital admission, multiple assessments during labour, and interventions during labour and childbirth. The alert line and action line, intrapartum monitoring parameters, and customized labour curves were assessed using the sensitivity, specificity, positive and negative likelihood ratios, diagnostic odds ratio and J statistic. **Outcomes:** Severe adverse birth outcomes. **Results:** The rate of severe adverse birth outcomes was 2.2% (223 women with severe adverse birth outcomes), the rate of augmentation of labour was 35.1% (3,506 women) and the caesarean section rate was 13.2% (1,323 women). 49% of women in labour crossed the alert line (4,163/8,489). All reference labour curves had a diagnostic odds ratio ranging from 1.29 to 1.60. The J statistic was less than 10% for all reference curves. **Conclusions:** Our findings suggest that labour is an extremely variable phenomenon and the assessment of cervical dilatation over time is a poor predictor of severe adverse birth outcomes. The validity of a partograph alert line based on the "one-centimetre per hour" rule should be re-evaluated..

Funding: Bill & Melinda Gates Foundation, United States Agency for International Development (USAID), UNDP/UNFPA/UNICEF/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction (HRP), and WHO (A65879).

Keywords: Partograph, Alert line, Childbirth, Diagnostic Accuracy, ROC space

Tweetable abstract: The alert line in check: results from a WHO study

Introduction

Labour and childbirth are natural processes with a relatively low frequency of complications among healthy pregnant women[1,2]. Intrapartum maternal and fetal monitoring is used to further minimize risks and is expected to enable early identification and prompt treatment of complications. The assessment of cervical dilatation is part of intrapartum monitoring and conducted by health care

providers to determine the adequacy of labour progress. The observed cervical dilatation is usually compared to reference labour curves to estimate the risk of labour complications and guide the use of interventions[3-7].

The WHO partograph is a decision-making support tool designed to assist health providers to identify women at risk of developing complications during labour and to guide the use of interventions intended to mitigate perceived risks[6]. With the partograph, the identification of certain patterns of cervical dilatation or other risk factors may prompt the transfer of the woman to higher-level health facility, intensification of intrapartum monitoring, and augmentation of labour or caesarean section[7]. The WHO partograph is based on clinical principles, including the notion that "normal" labour progress is defined by a cervical dilatation rate not less than one centimetre per hour between 4 and 10 centimetres of cervical dilatation[6]. This concept is the basis for the partograph "alert line" which was derived from average labour curves developed during the 1950's and 1960's[8-9]. Alternative and more recent labour curves have been developed to provide a reference for labour progress and serve as the basis for new partograph designs[5,10].

While some observational studies and other empirical evidence point towards benefit of using the WHO partograph, experimental, head-to-head comparisons failed to demonstrate an effect of the partograph in improving health outcomes related to labour and childbirth[11]. Furthermore, different studies have pointed to limitations in the "one-centimetre per hour rule" as a valid benchmark for assessing the adequacy of labour progress[5,10,12]. Our hypothesis is that if the partograph is unable to accurately identify women at risk of developing intrapartum complications, it will not be able to effectively guide labour management.

This manuscript reports on findings of the World Health Organization Better Outcomes in Labour Difficulty (BOLD) project. The present analysis assessed the diagnostic accuracy of the alert line, action line, and other parameters included in the WHO partograph as predictors of severe adverse birth outcomes. It also assessed the accuracy of customized labour curves to identify women at risk of developing severe adverse birth outcomes.

Methods

The BOLD project included quantitative, qualitative and service design research conducted in Nigeria and Uganda. Methodological details of the BOLD project have been described elsewhere[13,14]. This analysis is based on the quantitative component, a facility-based, multicentre, prospective cohort study. In brief, this study included women admitted for vaginal birth with single live foetuses during early first stage of labour across 13 hospitals in both countries. Women with spontaneous onset of labour presenting at cervical dilatation of ≤6 cm and those undergoing induction of labour took part in the study. Women with multiple pregnancies, gestational age less than 34 weeks 0 days, elective caesarean section, and those who were incapable of giving consent due to labour distress or obstetric emergencies at arrival were excluded. Participating institutions had a minimum of 1,000 deliveries per year with stable access to caesarean section, augmentation of labour, assisted vaginal birth. Midwives, obstetricians or obstetric residents provided intrapartum health care to women in labour. Doptones were used to assess fetal vital status at hospital admission and for intermittent monitoring through labour and childbirth. Labour management protocol, as well as the number and timing of pelvic examinations were not standardized across participating institutions. None of the institutions subscribed to the active management of labour protocol during the study period. Although the partograph was a standard element of medical records in all participating health facilities, its prospective application to guide labour management during the study period varied widely across the hospitals.

Eligible women were recruited into the study between December 2014 and November 2015. From the medical record, trained research nurses prospectively extracted detailed information on socio-demographic, anthropometric, obstetric and medical characteristics of study participants at hospital admission, multiple assessments during labour monitoring and interventions performed throughout the first and second stages of labour, and maternal and neonatal labour outcomes. Attending staff was approached to complement medical records data when needed. Data collection was limited to hospital stay of the mother and baby and there was no post-hospital discharge follow-up.

The current analysis was based on information on maternal baseline and admission characteristics, repeated assessments of cervical dilatation versus time, and maternal and neonatal outcome data. Severe adverse birth outcomes were defined as the occurrence of any of the following: stillbirths, intra-hospital early neonatal deaths, neonatal use of anticonvulsants, neonatal cardio-pulmonary resuscitation, Apgar score < 6 at 5 minutes, uterine rupture and maternal death or organ dysfunction with dystocia. Details of the sample size calculation are provided in the supporting information (Box S1).

Data analysis

Simple frequencies and proportions were used to describe the characteristics of the study population. Sensitivity, specificity, positive and negative likelihood ratios, diagnostic odds ratios and the J statistic (Youden's index) with 95% confidence intervals were used to estimate the diagnostic accuracy of the alert line and the action line in the identification of women who would develop a severe adverse birth outcome[15-18]. We used the true positive rate (i.e. sensitivity) and the false positive rate (i.e. 1-specificity) to graphically represent the diagnostic accuracy of the partograph parameters in the Receiver Operating Characteristic (ROC) space[19]. Each point estimate in the ROC space represents a classification result for binary parameters and the interpretation of the ROC space is similar to the ROC curve: optimal results are associated with high true positive rates combined with low false positive rates. The J statistic summarizes the performance of a binary classifier[16] and also expresses

the proportion of ideal performance of a diagnostic test (Box S2). The supporting information provides additional details related to the calculation and interpretation of these statistics (Tables S1, S2, S3).

The alert line and the action line are classifiers currently applied to all women regardless of their obstetric characteristics (e.g. nulliparous, multiparous, spontaneous or induced labour, previous caesarean section). We hypothesized that cervical dilatation curves customized according to the obstetric characteristics of the population could have a better accuracy than the generic alert and action lines. The study population was stratified into mutually exclusive, totally inclusive obstetric groups according to the 10-group Robson classification[20]: Group 1 (nulliparous, single cephalic pregnancy, 37 weeks gestation or more with spontaneous onset of labour), Group 2 (nulliparous women, single cephalic pregnancy, 37 weeks gestation or more, induced onset of labour), Group 3 (multiparous women without previous caesarean section, with single cephalic pregnancy, 37 weeks gestation or more, spontaneous onset of labour), Group 4 (multiparous women without previous caesarean section, with single cephalic pregnancy, 37 weeks gestation or more, induced onset of labour), Group 5 (All multiparous women with at least one previous caesarean section, single cephalic pregnancy, 37 weeks gestation or more), Group 10 (All women with singleton cephalic preterm pregnancy (less than 37 weeks gestation at childbirth)). Due to the eligibility criteria, this study has no woman from Group 8 (multiple pregnancies) or with caesarean section before labour. Women with non-cephalic presentations (Groups 6, 7 and 9) were grouped together. Groups 1-5 and 10, were further divided according to the use of augmentation of labour (present or absent), totalling 12 subgroups. Using data from women who did not have any severe adverse birth outcome, customized labour curves were generated for each of these 12 subgroups. Data from women pertaining to the groups 6, 7 and 9 were not used to generate customized curves due to small numbers. The customized cervical dilatation curves were created using a multi-state Markov model[21,22], which represented the cervical dilation pattern through intermediate states from 2 cm to 10 cm and

childbirth by selected percentiles and obstetric group (i.e. one labour curve for each obstetric group and selected percentile). In this model, each centimetre of cervical dilatation represented an intermediate state, and childbirth was the final, 'absorbing' state. The model was generated as a progressive unidirectional labour-to-childbirth model and the time of state change was determined by a set of transition intensities. The transition intensity represents the instantaneous likelihood of moving from one state to another and is generated as part of the multi-state Markov Model. For each one of the 12 obstetric subgroups, the multi-state Markov model generated labour curves representing labour progress of women faster or at the percentiles 50, 60, 70, 80, 90 and 95.

Once the percentile curves were generated for each obstetric subgroup of women without severe adverse birth outcomes, women were classified as having crossed or not having crossed each of the percentile curves of their relevant obstetric subgroup. The study population was then consolidated and all women who crossed their relevant percentile 50 curves were grouped together (i.e. women slower than the customized percentile 50). Similarly, women were classified as being slower or faster/equal to the relevant percentile 60, 70, 80, 90 and 95. We estimated the accuracies of the customized percentiles curves in the identification of women who would develop a severe adverse birth outcome by comparing women with labour progress slower than the specific percentile to those faster or equal to that percentile. Sensitivity, specificity, positive and negative likelihood ratios, diagnostic odds ratios with 95% confidence intervals, J statistic and ROC space plotting were used to estimate the accuracy of the percentile curves in the identification of women who would develop a severe adverse birth outcome.

Statistical analysis were carried in R language[23] and Microsoft Excel (2010).

Funding: Bill & Melinda Gates Foundation, United States Agency for International Development (USAID), UNDP/UNFPA/UNICEF/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction (HRP), and WHO (A65879). The external funders played no role in conducting research and writing this article.

Results

The analysis flow is shown in Figure 1. Thirteen hospitals (nine from Nigeria and four from Uganda) and a total of 9,995 women (4,964 from Nigeria and 5,031 from Uganda) took part in this study. The average age of the participants was 27.9 years (± 5.0); 3.2% were less than 20 years (320 women) and 11.0% were 35 or more years old (1,100 women). The majority of the participants had a partner (97.6%, 9,753 women); 5.2% of the participants had either incomplete primary education or no education (525 women), 5.6% had complete primary education (564 women), 42.5% had either complete or incomplete secondary education (4,245 women) and 45.4% had either complete or incomplete post-secondary/tertiary education (4,537 women). A total of 4,076 nulliparous women took part in the study (40.8%) and among women with at least one previous birth (59.2%, 5,919 women), 535 (5.4%) had a previous caesarean section. A total of 667 women (6.7%) had no antenatal care visit, 4229 (42.3%) had one to three antenatal care visits and 5,007 (50.6%) had four visits of more; 1228 women (12.3%) developed pre-labour complications during the current pregnancy. The majority of women initiated labour spontaneously (8,984 women, 89.9%) at a gestational age between 37-41 weeks (91.2%, 9111 women) with only 594 (5.9%) being referred in labour from another health facility. All women participating in this study had singleton pregnancies with 98.6% of them (9,845 women) with a cephalic presentation. The mean number of cervical assessments between 4 and 10 cm was 2.22 (± 1.02). Augmentation of labour was used in 3,506 women (35.1%). Pharmacological analgesia was rarely used (2.0%, 196 women). Table 1 presents the distribution of the study population according to the 10-group Robson classification. The overall intrapartum caesarean section rate was 13.2% (1,323 women) and the rate of severe adverse birth outcomes was 2.2% (223 women with severe adverse birth outcomes, Table S4). Nearly half of women with at least two assessments of cervical dilatation between 4 and childbirth crossed the alert line (49%; 4,163/8,489). Figure 2 illustrates the progress of labour of the study population: on the top image, each gray line represents the progress of individual women without severe adverse birth outcomes, each red line represents the progress of individual women with adverse outcomes; on the bottom

image, labour curves for 95th percentile of women in labour without augmentation of labour is displayed by obstetric group. Video S1displays an animation of labour progress of all women in labour that reached at least four centimetres of cervical dilatation.

The sensitivity, specificity, positive and negative likelihood ratios, diagnostic odds ratio and J statistic for the WHO partograph alert and action lines and the percentiles 50, 60, 70, 80, 90 and 95 are presented in Tables 1 and Table S5. Across all reference curves, women who crossed them tended to show a mild increase in the odds of severe adverse birth outcomes when compared to those who did not cross the reference line. All reference curves had a diagnostic odds ratio ranging from 1.29 to 1.60. All reference curves had positive likelihood ratios smaller than 1.5 and negative likelihood ratios greater than 0.85. The J statistic was less than 10% for all reference curves. Figure 3 presents the ROC space analysis, with all the above-mentioned predictors showing a poor diagnostic performance.

Table S6 and Figure S1 present the diagnostic accuracy of various predictors included in the partograph (the definitions of these predictors are presented in Table S7). Abnormal fetal heart rate, absence of fetal movements, significant moulding, significant caput succedaneum, meconium, and maternal hyperthermia (fever) were associated with mild to moderate increased odds of severe adverse birth outcomes. Similarly to the labour curves, the examined predictors presented poor performance for the prediction of severe adverse birth outcomes.

Discussion

Main findings

Labour is an extremely variable phenomenon and our findings suggest that the assessment of cervical dilatation over time is a poor predictor of severe adverse birth outcomes. Labour curves depicting the cervical dilatation over time (including the WHO Partograph Alert and Action Lines) showed poor diagnostic accuracy to identify women at risk of severe adverse birth outcomes during labour. We

draw one main inference from these findings: the validity of a partograph alert line based on the "one-centimetre per hour" rule should be re-evaluated.

Strengths and Limitations

These findings are relevant to care provided in health facilities, particularly in Sub-Saharan Africa, and have potential implications for clinical practice. However, despite the procedures adopted to ensure appropriate study implementation and high quality data, some limitations need to be considered. The primary data source in this study was the routine hospital record complemented by information obtained from clinical staff. We opted for this approach to minimize interference with the standard practice in health facilities, but acknowledge that it could be associated with irregular and, at times, incomplete intermittent assessment and record of maternal and fetal status during labour. Although unlikely due to clinical workload, availability of Doptones provided by the study in the labour wards may have facilitated fetal monitoring and contributed to increased identification of fetal distress, which could have impacted in the clinical management and outcomes. We were also able to determine the fetal vital status at arrival of all women, which resulted in accurate assessment of intra-partum, intra-hospital fetal mortality. This assessment enabled disentangling pre-hospital fetal deaths from intra-hospital fetal deaths and uncovered a low intra-hospital fetal mortality, despite the constraints to optimal care in health facilities. None of the participating hospitals subscribed to a systematic implementation of active management of labour; while this could contribute to a less standardized management of labour it favoured a less interventionist approach and an intra-hospital labour progression that was more closely related to the natural progression in many women. Given the differences of workload and health facility protocols, standardization of intrapartum maternal-fetal monitoring and record was a challenging task. Several mechanisms were used to minimize methodological heterogeneity and increase data quality as much a possible (such as research assistant training, use of a visual check of the data collection forms before data entry, automated queries, double-checking of selected medical records, and thorough audit of unclear cases, especially those resulting in mortality). It should also be considered that crossing the alert or action lines could

have prompted health providers to implement interventions in the current cohort population. These interventions could have modified the final outcome either for good or bad.

Interpretation

Health facilities in low-resource settings often struggle with a shortage of human resources and life-saving commodities, training resources, and health infrastructure, which limit early identification and effective management of labour complications. Conversely, overestimation of risk of complications and over-medicalization of care during labour and childbirth may lead to iatrogenic complications, avoidable suffering and waste of limited resources[24]. In an attempt to optimize intrapartum care, several organizations recommend the use of the WHO partograph to guide labour monitoring and management[24]. The "one-centimetre per hour rule", as illustrated by the partograph alert line, has also (formally or informally) been used to prompt labour interventions in many settings around the world[4,25]. Global efforts to promote the use of the partograph in the last three decades have been met with mixed results. While most health care providers working in maternity settings know the partograph, it is frequently used retrospectively for recording purposes instead of providing prospective support to clinical decision-making. Possible reasons for these shortcomings include difficulties in its use and interpretation[26,27]. Our findings suggest that the poor predictive performance of the partograph – and the consequent effect in supporting effective decision making –could contribute to the lack of interest in using the tool prospectively.

As countries navigate through the obstetric transition[28], a marked trend toward medicalization of labour and childbirth is observed. Several determinants of medicalization of labour and childbirth are at play, including models of care based on the notion that a normal labour abides by the "one-centimetre per hour" rule. This notion has been embedded in generations of health care providers across the world and the implicit or explicit influence of this notion in the obstetric and midwifery culture cannot be over-emphasized. However, as suggested by our findings, a cervical dilatation rate of "one-centimetre per hour" may be unrealistically fast for a substantial proportion of women in

labour. The mismatch between the unrealistic expectations of health care providers and the physiology of labour may give rise to the constructed "need" of an intervention in a natural process that could otherwise be slower than currently expected but end well and naturally. The poor accuracy of the tool means in one hand that a high proportion of women would receive an intervention without a valid justification and on the other hand women at risk would not be recognized in time to avoid the adverse outcome. The excessive use of interventions may also contribute to adverse outcomes. For example, augmentation of labour is a well-established risk factor for fetal distress; unnecessary augmentation of labour, prompted by the "one-centimetre per hour" rule, may be harmful particularly in settings with limited capacity for providing appropriate, intermittent fetal monitoring. Another potential adverse effect of the above mismatch is increased tension, anxiety and frustration among the staff, which could be a contributing factor to disrespect, abuse and mistreatment of women during labour and childbirth. However, allowing an increase in the average duration of labour in health facilities has a direct impact in the occupancy rate of labour-ward beds which could further complicate shortage of hospital beds and overcrowding of health facilities. Reducing the amount of interventions during labour could reduce staff workload. Research to determine the short and long-term consequences of a less-invasive intrapartum care model at the individual and at the health systems level is warranted.

The poor performance of the partograph and the customized labour curves may not be a surprising finding. The rationale for using the partograph for preventing labour problems goes back several decades when it was introduced for timely referral from peripheral health facilities to prevent complications of obstructed labour[9]. Fetal and early neonatal outcomes are much more likely to be impacted by events that are not related to cervical dilatation rate such as placental abruption, cord compression, cord prolapse, meconium aspiration, intra-uterine growth restriction among many other reasons. In South Africa, for example, only 6% of fetal and early neonatal deaths were associated with prolonged labour[29]. However, we should not overlook the finding that slower labours compared to faster labours (in different percentiles) tended to be associated with a mild increase in

the risk of adverse outcomes. Nevertheless, this association alone can hardly provide a basis for a reliable classification tool due to the excessive number of false positives. For example, nearly half of the study population crossed the alert line, making the policy of transferring women who crossed the alert line to referral hospitals impractical. In this context, and given the limitations of static, paper-based diagnostic tools, the development and testing of more sophisticated, dynamic, easy-to-use tools for improved risk classification during labour is a priority. A cluster-randomized trial, comparing a static paper-based partograph to a dynamic, multivariable prediction model would be the ideal research to be carried out next.

Conclusion

Our findings suggest that the validity of a partograph alert line based on the "one-centimetre per hour" rule should be re-evaluated. Labour is an extremely variable phenomenon and emphasis should be given to individualized, supportive, person-centered care during labour and childbirth.

Disclosure of interest

We declare that we have no conflicts of interest. Completed ICMJE disclosure forms are available to view online as supporting information.

Contribution to authorship

JPS, OTO and AMG developed the research protocol with input from members of the project steering committee and advisory group, Nigeria and Uganda research teams, data science team, and the project coordination and support team. The analysis plan was developed by JPS with input from OTO. Data analysis was carried out by RR, FBJ, LOC and JPS. JPS drafted this manuscript with substantial contributions from OTO, BF, KM, RR, FBJ, LOC, DA, AMG. All authors reviewed the draft manuscript for intellectual content and approved the final manuscript for publication.

Details of ethics approval

Scientific and technical approval was obtained from the Review Panel on Research Projects (RP2) of UNDP/UNFPA/UNICEF/WHO/World Bank Special Program of Research, Development and Research Training in Human Reproduction (HRP). Ethical approval was obtained from the World Health Organization Ethical Review Committee (protocol A65879, approval date 25 August 2014), the Makerere University School of Health Sciences Research and Ethics Committee, Uganda (protocol #SHSREC REF 2014-058), University of Ibadan/University College Hospital Ethics Committee (UI/EC/14/0223), Federal Capital Territory Health Research Ethics Committee, Nigeria (protocol FHREC/2014/01/42/27-08-14) and Ondo State Government Ministry of Health Research Ethics Review Committee, Nigeria (AD 4693/160).

Acknowledgement

This manuscript was developed as part of the World Health Organization's BOLD ("Better Outcomes in Labour Difficulty") project. The BOLD project implementation was a collaborative effort of a large number of academic staff, hospital personnel and researchers from 13 hospitals across Uganda and Nigeria. We acknowledge the contributions of all midwives, doctors, facility administrators, and all maternity unit staff in the maternity units of all hospitals that participated in the project.

The authors would like to thank the members of the following groups for their input during project implementation and analysis:

BOLD project steering committee & advisory group: (Priya Agrawal, Fernando Althabe, Deborah Armbruster, France Donnay, G Justus Hofmeyr, Vanora Hundley, Jen Jardine, David Kyaddondo, Tina Lavender, Simon Lewin, Jerker Liljestrand, Matthews Mathai, Lanre Olutayo, Janna Patterson, Robert Pattinson, Kate Ramsey, Harshadkumar C Sanghvi, Mary Ellen Stanton, Petra Tenhoope-Bender, Mari Tikkanen, Melanie Wendland, Jim Zhang), Nigeria research team (Amos Adebayo, Nike Adesina, Adesina Akintan, Olubunmi Alabi, Francis E Alu, Halima Bello, Felicia Fasipe, Bukola Fawole, Hadiza Idris, Oluwatosin Lawal, Ola Okike, Lawal Oyeneyin, Sanni Wilfred), Uganda research team (Salim Bataale, Michael Bukennya, Josaphat Byamugisha, Wasswa Damien, Kidza Mugerwa, Othiniel Musana, Miriam Nakalembe), data science team (Domingos Alves, Francisco Barbosa-Júnior, Gleise Kele Bernucci, Hayala Cristina Cavenague, Livia Ciabati, Alexandre Cristovão Maiorano, Jéssica Marcelino, Leonardo Moscovici, Lariza Laura de Oliveira, Gleici Perdoná, Rodrigo Reis, Tiago Lara Michelin Sanches, Joao Paulo Souza), and the project coordination and support team (Lily Zilla Atutornu, Rajiv

Bahl, Meghan Bohren, Metin Gülmezoglu, Catherine Cécile Kiener, Olufemi T Oladapo, Joao Paulo Souza, Özge Tunçalp, Joshua P Vogel). The manuscript presents the views of the named authors only.

Funding: Bill & Melinda Gates Foundation, United States Agency for International Development (USAID), UNDP/UNFPA/UNICEF/WHO/World Bank Special Programme of Research, Development and Research Training in Human Reproduction (HRP), and WHO (A65879). The external funders played no role in conducting research and writing this article.

References

1. Snowden JM, Tilden EL, Snyder J, Quigley B, Caughey AB, Cheng YW. Planned Out-of-Hospital Birth and Birth Outcomes. N Engl J Med. 2015 Dec 31;373(27):2642-53. doi: 10.1056/NEJMsa1501738.
2. de Jonge A, Geerts CC, van der Goes BY, Mol BW, Buitendijk SE, Nijhuis JG. Perinatal mortality and morbidity up to 28 days after birth among 743 070 low-risk planned home and hospital births: a cohort study based on three merged national perinatal databases. BJOG. 2015 Apr;122(5):720-8. doi: 10.1111/1471-0528.13084.
3. Cohen WR, Friedman EA. The assessment of labor: a brief history. J Perinat Med. 2017 Mar 22. pii: /j/jpme.ahead-of-print/jpm-2017-0018/jpm-2017-0018.xml.doi: 10.1515/jpm-2017-0018. [Epub ahead of print] PubMed PMID: 28328535.
4. O'Driscoll K, Stronge JM. Active management of labor. British medical journal 1973; 3(5880): 590.

5. Zhang J, Landy HJ, Branch DW, Burkman R, Haberman S, Gregory KD, et al. Contemporary patterns of spontaneous labor with normal neonatal outcomes. Obstetrics and gynecology 2010; 116(6): 1281-7.
6. World Health Organization (1993a) Preventing prolonged labour: A Practical Guide. The Partograph Part I. Principle and Strategy. WHO, Geneva WHO/FHE/MSM 93.8
7. World Health Organization (2000) Managing complications in pregnancy and childbirth: a guide for midwives and doctors. IMPAC. WHO/RHR.007. World Health Organization, Geneva
8. Friedman E. The graphic analysis of labor. American journal of obstetrics and gynecology 1954; 68(6): 1568-75.
9. Philpott RH, Castle WM. Cervicographs in the management of labour in primigravidae. I. The alert line for detecting abnormal labour. J Obstet Gynaecol Br Commonw. 1972 Jul;79(7):592-8. PubMed PMID: 5043422.
10. Neal JL, Lowe NK. Physiologic partograph to improve birth safety and outcomes among low-risk, nulliparous women with spontaneous labor onset. Med Hypotheses. 2012 Feb;78(2):319-26. doi: 10.1016/j.mehy.2011.11.012. Epub 2011 Dec 3. PubMed PMID: 22138426; PubMed Central PMCID: PMC3254242.
11. Lavender T, Hart A, Smyth RM. Effect of partogram use on outcomes for women in spontaneous labour at term. Cochrane Database Syst Rev. 2013 Jul 10;(7):CD005461.
12. Hamilton EF, Warrick PA, Collins K, Smith S, Garite TJ. Assessing first-stage labor progression and its relationship to complications. Am J Obstet Gynecol. 2016 Mar;214(3):358.e1-8.

13. Souza JP, Oladapo OT, Bohren MA, Mugerwa K, Fawole B, Moscovici L, et al. The development of a Simplified, Effective, Labour Monitoring-to-Action (SELMA) tool for Better Outcomes in Labour Difficulty (BOLD): study protocol. Reproductive health 2015; 12: 49.

14. Bohren MA, Oladapo OT, Tunçalp Ö, Wendland M, Vogel JP, Tikkanen M, et. al.. Formative research and development of innovative tools for "Better Outcomes in Labour Difficulty" (BOLD): study protocol. Reprod Health. 2015 May 26;12:50.

15. Powers DM. Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation. Journal of Machine Learning Technologies. 2011, 2(1): 37–63.

16. Youden WJ. Index for rating diagnostic tests. Cancer. 1950 Jan;3(1):32-5.

17. Glas AS, Lijmer JG, Prins MH, Bonsel GJ, Bossuyt PM. The diagnostic odds ratio: a single indicator of test performance. J Clin Epidemiol. 2003 Nov;56(11):1129-35.

18. Cook C, Cleland J, Huijbregts P. Creation and Critique of Studies of Diagnostic Accuracy: Use of the STARD and QUADAS Methodological Quality Assessment Tools. J Man Manip Ther. 2007;15(2):93-102.

19. Fawcett T. An Introduction to ROC Analysis. Pattern Recognition Letters. 2006, 27(8): 861–874.

20. Robson MS. Classification of caesarean section. Fetal and Maternal Medicine Review.2001;12:23–39. doi: 10.1017/S0965539501000122.

21. Jackson CH. Multi-state models for panel data: The msm package for R. Journal of

Statistical Software.2011;38(8):1–29.

22. Reis R. Labour Curves and Markov models. MSc [Dissertation]. Ribeirao Preto (Brazil): University of Sao Paulo; 2017

23. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2014 (http://www.R-project.org)

24. Miller S, Abalos E, Chamillard M, Ciapponi A, Colaci D, Comandé D, et al.Beyond too little, too late and too much, too soon: a pathway towards evidence-based, respectful maternity care worldwide. Lancet. 2016 Oct 29;388(10056):2176-2192.

25. Maffi I. The Detour of an Obstetric Technology: Active Management of Labor Across Cultures. Med Anthropol. 2016;35(1):17-30.

26. Bedwell C, Levin K, Pett C, Lavender DT. A realist review of the partograph: when and how does it work for labour monitoring? BMC Pregnancy Childbirth. 2017 Jan 13;17(1):31.

27. Ollerhead E, Osrin D. Barriers to and incentives for achieving partograph use in obstetric practice in low- and middle-income countries: a systematic review. BMC Pregnancy Childbirth. 2014 Aug 16;14:281. doi: 10.1186/1471-2393-14-281.

28. Souza JP, Tunçalp Ö, Vogel JP, Bohren M, Widmer M, Oladapo OT, et al. Obstetric transition: the pathway towards ending preventable maternal deaths. BJOG. 2014 Mar;121 Suppl 1:1-4.

29. Saving babies 2012-2013: Ninth report on perinatal care in South Africa. RC

Pattinson, N Rhoda, Tshepesa Press, Pretoria, 2014

30. Grimes DA, Schulz KF. Refining clinical diagnosis with likelihood ratios. Lancet. 2005 Apr 23-29;365(9469):1500-5.

Table 1: Diagnostic accuracy of the alert line and the action line for severe adverse birth outcomes (N=8603)

Clinical Signs		**Severe adverse birth outcomes**		**Sensitivity**	**Specificity**	**Positive Likelihood Ratio**	**Negative Likelihood Ratio**	**Diagnostic Odds Ratio**	**J statistic**
		Present	**Absent**	**(95% CI)**	**(95% CI)**	**(95% CI)**	**(95% CI)**	**(95% CI)**	**[Youden's Index] (95% CI)**
Alert line	Crossed	110	4053	56.7% (49.7-63.5)	51.1% (50.1-52.2%)	1.16 (1.02-1.32)	0.85 (0.72-1.00)	1.37 (1.03-1.83)	7.8% (0.8-14.9)
	Not crossed	84	4242						
Action line	Crossed	38	1230	19.6% (14.6-25.7)	85.2% (84.4-85.9)	1.32 (0.99-1.77)	0.94 (0.88-1.01)	1.40 (0.98-2.01)	4.8% (-0.9 – 10.4)
	Not crossed	156	7065						

TP: True Positive; TN: True Negative; FP: False Positive; FN: False Negative; Sens: sensitivity; Sp: specificity

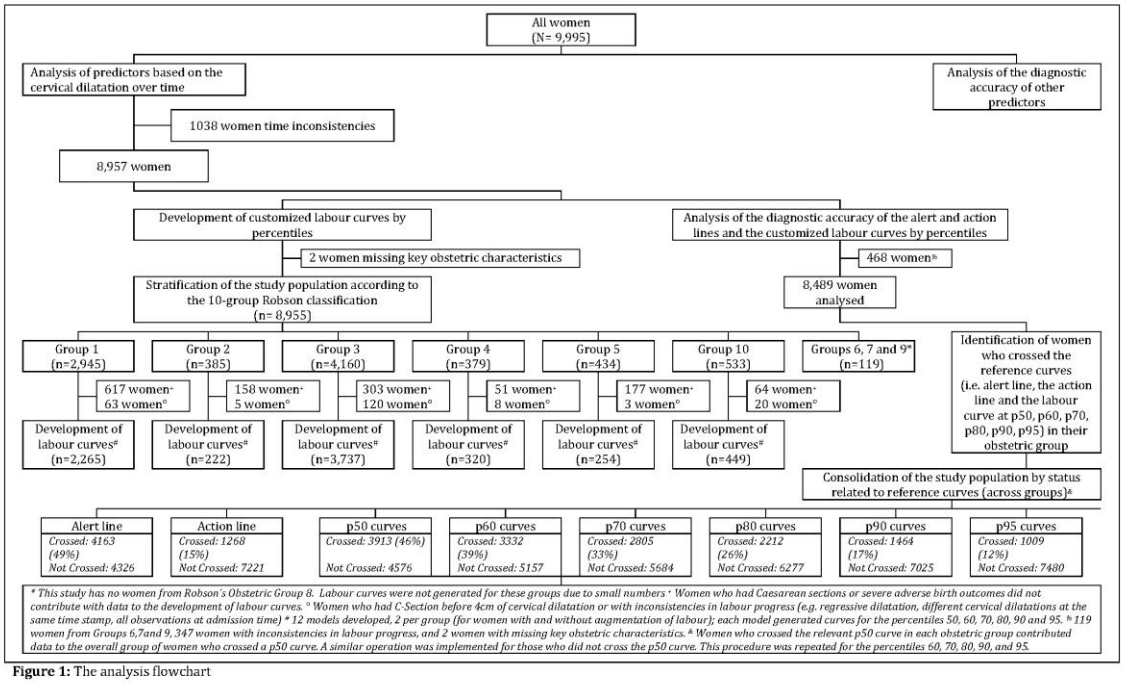

Figure 1: The analysis flowchart

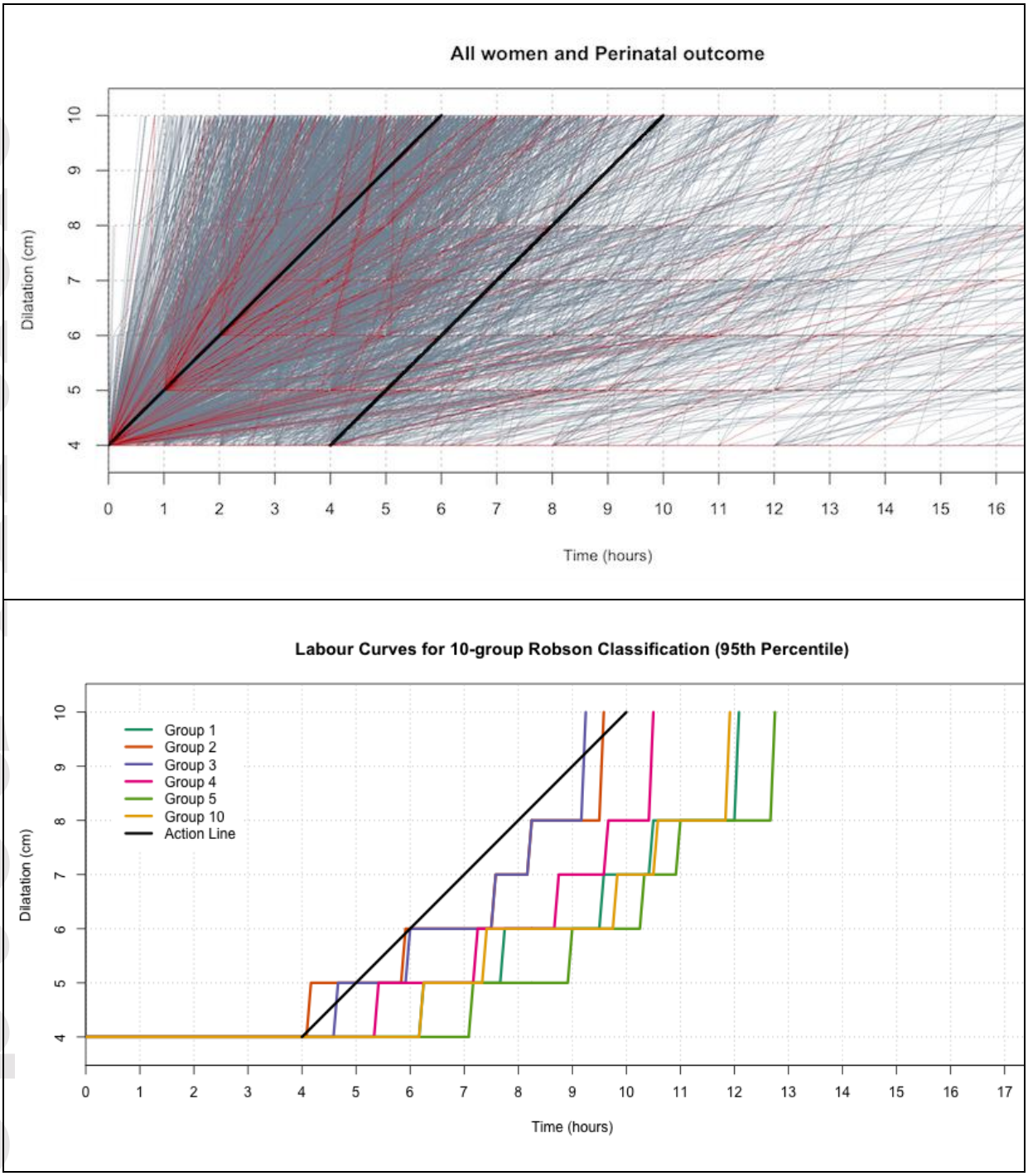

Figure 2: *(Top)* Cervical dilatation over time (all women with at least 2 cervical dilatation assessments between four centimetres and childbirth). Grey lines denote labour progress of women without severe adverse birth outcomes; red lines denote labour progress of women with severe adverse birth outcomes. *(Bottom)* Labour curves for selected groups of the 10-Group Robson Classification (95th Percentile, women without augmentation of labour)

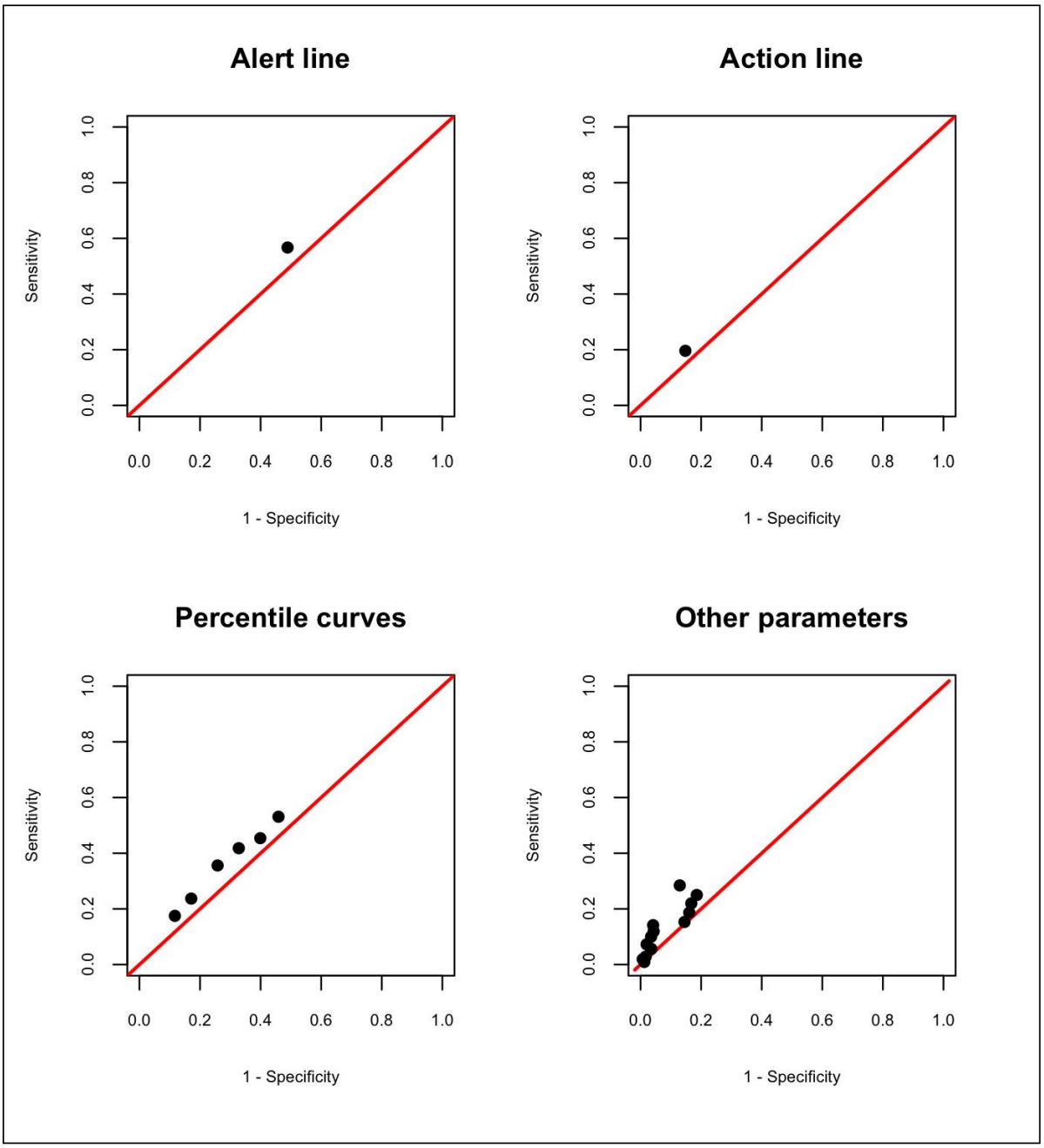

Figure 3: Analysis of the ROC space (alert and action line, customized percentile curves, and other parameters included in the partograph)