Newborn skin maturity models for gestational age prediction

Zilma Silveira Nogueira Reis (zilma@ufmg.br)
Department of Gynecology and Obstetrics, Faculty of Medicine, Universidade Federal de Minas Gerais, Brazil. Center for Artificial Intelligence, Innovation and Health, Universidade Federal de Minas Gerais, Brazil. https://orcid.org/0000-0001-6374-9295

Rodney Nascimento Guimarães
Department of Gynecology and Obstetrics, Faculty of Medicine, Universidade Federal de Minas Gerais, Brazil. https://orcid.org/0000-0001-5617-4470

Roberta Maia de Castro Romanelli
Department of Pediatrics, Faculty of Medicine, Universidade Federal de Minas Gerais, Brazil. https://orcid.org/0000-0002-1660-0751

Juliano de Souza Gaspar
Department of Gynecology and Obstetrics, Faculty of Medicine, Universidade Federal de Minas Gerais, Brazil. https://orcid.org/0000-0003-0670-9021

Gabriela Silveira Neves
Hospital Sofia Feldman, Brazil. Department of Pediatrics, Faculty of Medicine, Universidade Federal de Minas Gerais, Brazil. https://orcid.org/0000-0001-6765-2968

Marynea Silva do Vale
Maternal and Child Unit, University Hospital, Universidade Federal do Maranhão, EBSERH, Brazil. https://orcid.org/0000-0002-1544-5470

Paulo de Jesus H Nader
University Hospital of Canoas, Universidade Luterana do Brasil, Brazil.

Martha David Rocha de Moura
Hospital Materno Infantil de Brasília, Brazil. https://orcid.org/0000-0002-7851-375X

Gabriela Luíza Nogueira Vitral
Department of Pediatrics, Faculty of Medicine, Universidade Federal de Minas Gerais, Brazil. https://orcid.org/0000-0001-7306-8776

Marconi Augusto Aguiar dos Reis
Department of Pediatrics, Faculty of Medicine, Universidade Federal de Minas Gerais, Brazil. https://orcid.org/0000-0001-5294-1361

Marcia Margarida Mendonça Pereira
Maternal and Child Unit, University Hospital, Universidade Federal do Maranhão, EBSERH, Brazil. https://orcid.org/0000-0002-2699-2963

Patrícia Franco Marques

Maternal and Child Unit, University Hospital, Universidade Federal do Maranhão, EBSERH, Brazil. https://orcid.org/0000-0001-6812-6687

Silvana Salgado Nader

University Hospital of Canoas, Universidade Luterana do Brasil, Brazil. https://orcid.org/0000-0001-9360-6201

Augusta Luize Harff

University Hospital of Canoas, Universidade Luterana do Brasil, Brazil. https://orcid.org/0000-0003-3220-1643

Ludmylla de Oliveira Beleza

Hospital Materno Infantil de Brasília, Brazil. https://orcid.org/0000-0001-9975-562X

Maria Eduarda Canellas de Castro

Hospital Materno Infantil de Brasília, Brazil. https://orcid.org/0000-0001-8727-8328

Rayner Guilherme Souza

Department of Gynecology and Obstetrics, Faculty of Medicine, Universidade Federal de Minas Gerais, Brazil. https://orcid.org/0000-0003-2440-2242

Gisele Lobo Pappa

Center for Artificial Intelligence, Innovation and Health, Universidade Federal de Minas Gerais, Brazil. Computer Science Department, Universidade Federal de Minas Gerais, Brazil. https://orcid.org/0000-0002-0349-4494

Regina Amélia Pessoa Lopes de Aguiar

State Health Department of Minas Gerais, Brazil https://orcid.org/0000-0003-2470-3539

Research Article

Keywords: Gestational Age, Infant, Premature, Skin Physiological Phenomena, Photomedicine, Equipment and Supplies

Posted Date: January 7th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1216628/v2

Abstract

A multicenter clinical trial evaluated the accuracy of a novel device to detect preterm newborns. A portable multiband reflectance photometric device assessed 781 newborns' skin maturity and used machine learning models to predict reference gestational age, adjusting it to birth weight and antenatal corticosteroid therapy exposure. The day difference between the reference and the test had a median of -1.4 (IQR: -2.1). Using established methods such as comparator ultrasound and last menstrual period (LMP), the medians were 0 (IQR: 4) and 0.01 (IQR: 4), respectively. For prematurity discrimination, the area under the receiver operating characteristic curve (AUROC) was 0.986 (95% CI: 0.977 to 0.994). In newborns with absent or unreliable LMP, the intent-to-discriminate analysis showed that the test generated correct classifications 95.8% of the time. The assessment of the newborn's skin maturity adjusted by learning models promises accurate pregnancy dating at birth without the use of antenatal ultrasound reference.

Introduction

Public policies and best practices advise pregnant women to plan pregnancy with early access to prenatal care for pregnancies to be safely monitored until birth.[1] However, many barriers to covering all pregnancies and births with due care remain unsurpassed, particularly in scenarios without well-equipped facilities.[2,3] Adverse maternal and neonatal outcomes affect newborns unevenly according to the birth scenario and gestational age.[4] First-day-of-life mortality is 30 times higher in low- and medium-income countries than in high-income countries.[5] Preterm neonates are more vulnerable to death or survival with neurological sequelae. The need to recognize early risks at birth is faced with reduced early access to prenatal care and high-cost technologies for pregnancy dating, such as obstetric echography in resource-constrained settings.[6,7] Early obstetric ultrasound currently offers the best gestational age method.[8] However, access to high-cost equipment, poor training, lack of skills of health professionals, or late prenatal care limit pregnancy dating and, consequently, detection of prematurity.[7,9] Improving preterm birth outcomes demands accurate gestational age assessment to direct opportune decisions regarding neonatal care.[10,11] Approaches to enhance the reliability of pregnancy dating through more accurate and accessible technologies can improve pregnancy outcomes and neonatal survival.[8,12]

Health technology development is critical for supporting healthcare systems. Medical devices and digital health technologies have brought innovative solutions with the potential to save lives,[13] mitigating gaps of quality among disparate healthcare scenarios.[14] Furthermore, digital health technologies emerge with the potential to impact the equality of healthcare, creating new landscapes of opportunities such as applied data science to improve prediction models.[15,16] Currently, computer science has advanced with improvements to medical practice, detecting patterns by processing datasets through layered mathematical models,[17] fostering skills and competencies of professionals to support the best healthcare decisions.[18]

The new test explored in this study is an innovative approach used to estimate gestational age based on the photobiological properties of the newborn's skin and using learning predictive models enhanced with clinical variables.[19] Usable as a medical device, we developed this technology to easily assist health professionals in making better decisions regarding newborn care whenever the pregnancy dating is unknown or doubtful. This study aims to validate the photobiological model of skin maturity adjusted to clinical data to promptly detect gestational age and determine its accuracy in detecting prematurity. We tested the hypothesis of equivalence between the gestational age measured by this new test, the pregnancy dating comparators calculated using ultrasound exams, and the last menstrual period (LMP).

Results

Of the 791 potential eligible newborns, two were under Rh alloimmunization during pregnancy, which was considered an exclusion criterion (Fig. 1). Among the 789 newborns with their skin assessed with the medical device, eight had no reference standard to allow the dependent variable; four had no mandatory embryo measurement; three had no second ultrasound comparator; and one had an unsolved digit date error. All 781 newborns who met the eligibility criteria of the clinical trial were included in the analysis. According to the set of models, two predictive algorithms provided gestational age values, as detailed in the Gestational age prediction by the new test subsection.

Seven hundred four women gave birth to 781 newborns. Despite early access to prenatal care (median 9, IQR 4 weeks of gestation), only 296 (48.2%) women fulfilled the criteria for valid LMP, among 614 who were able to provide such a date (Table 1). A number of the pregnant women (n= 215, 30.7%) received antenatal corticosteroid therapy for fetal maturation (ACTFM), following local protocols, corresponding to 273 / 777 (35.1%) newborns. According to reference gestational age at birth, 415 (53.1%) newborns were term. Among 366 (46.9%) preterm newborns, 235 (30.1%) had a gestational age at birth less than 37 to 32 weeks, and 131 (16.8%) had a gestational age less than 32 to 24 weeks, achieving approximately the expected arranged proportion for the sampling.[20] The frequency of fetal abnormal growth classification at birth was 116 (14.9%) who were small for gestational age and 59 (7.6%) who were large for gestational age.

Gestational age using established methods of pregnancy dating

The distribution of gestational age calculated according to the different bases, reference (R), a second ultrasound exam after 13 weeks and 6 days of gestation and before 22 weeks (C1), and LMP (C2) corroborated the existing differences among established methods of antenatal dating figured in the overlapped histogram (Fig. 2). There were 27 (3.5%) posterm births when using comparator C2, 2 (0.3%) when using comparator C1, and 1 (0.1%) when using standard R for gestational age dating. Concerning the data quality of medical history reference used in gestational age estimation, the analysis of day preference confirmed bias for the LMP reference, the analysis of which is in Supplementary Note 1.

Correlation between the skin reflectance of the newborn and the reference gestational age

Newborn skin reflection acquisition had a positive high correlation with the reference standard gestational age, Pearson correlation coefficient = 0.79, R^2 = 0.62, p < 0.001 (Fig. 3).

Gestational age prediction by the new test

We considered the best parsimonious models according to available variables in the birth scenario. When the birth weight was known and ACTFM exposure information was not, Model 1 fitted gestational age with a linear-ridge algorithm. In cases when the birth weight and ACTFM information were known, Model 2 fitted gestational age with an XGBoost algorithm. The set of models adjusted with machine learning models is presented in Supplementary Table in Note 2.

The boxplots in Fig. 4a present gestational ages calculated by applying the predictive algorithms: R-gestational age and the comparators C1- and C2- gestational age, highlighting preterm birth rates. Data concerning the gestational age distribution of frequency according to the methods are in Table 2. The agreement between Model 1, Model 2, and the R-gestational age was high, as well when analyzing predicted gestational ages against the comparators C1- and C2- gestational age. The Bland-Altman 95% limits were -28.6 to 28.6 days for Model 1 and -19.6 to 19.6 days for Model 2 (gestational age), Table 3.

With a focus on the day difference between methods of determining R-gestational age, Fig. 4b displays box plots for each method of estimation. The day difference had a median of -1.4 (IQR: -2.1) for the Model 2-gestational age, and a median of 0.21 (IQR: 14) days for the Model 1 gestational age (Fig. 4b and Table 2). Established methods of pregnancy dating (C1 and C2) had median differences of 0 (IQR: 4) and -0.01 (IQR: 4), respectively (Table 2). Moreover, the proportion of preterm newborns correctly detected at birth within a one-week error is presented as the value within the interval between hatched lines (Fig. 4b). We noticed that Model 2-gestational age achieved 98.7% agreement with the reference pregnancy dating within one-week error, and Model 1-gestational age had 53.5%. Thus, the ACTFM exposure information added advantages to the skin reflectance and birth-weight-based Model 2.

Accuracy of the new test to discriminate preterm newborns

Skin reflectance itself had good accuracy in discriminating preterm term newborns, with AUROCs of 0.956 (95% CI: 0.942, 0.970) at 37 weeks, 0.937 (95% CI: 0.916, 0.958) at 32 weeks, and 0.959 (95% CI: 0.936, 0.981) at 28 weeks (cutoff) (Supplementary Fig. 1). The overall accuracy of preterm newborn classification is presented in Fig. 5 according to the predictive Model 1- and Model 2-gestational age considering different prematurity cutoff points. Regarding the 95% confidence interval (CI) analysis for AUROC, Model 2 gestational age had higher performance in discriminating preterm against term newborns at 37 weeks than Model 1 gestational age (Fig. 5). For 32- and 28-week cutoffs, Model 1- and Model 2-gestational age had similar AUROCs, considering an overlapping of 95% CIs. A comprehensive analysis of prediction accuracy for preterm newborns by the methods of gestational age estimation for different prematurity cutoffs is shown in Table 3.

Intent to preterm newborn discrimination by the new test

Birth care settings where the new technology is to be applied deserve an intent to preterm newborn discriminant analysis, simulating the existence of baseline references for the gestational age calculation. Therefore, we considered the lack of LMP recall or unreliable information as scenario one, corresponding to 57.7% of the sample (Table 4). The lack of a reliable LMP resulted in low accuracy (69.6%) of discrimination with imprecise dates. Nevertheless, 86.9% of the newborns were correctly classified as preterm or term using Model 1-gestational age. Using gestational age based on Model 2, 95.8% of the newborns were correctly classified as preterm or term, with a 0.5% false-positive rate. Great accuracy using any available method for gestational age estimation was observed in scenario two, where the LMP was reliable.

Safety of the device

There were no reports of unexpected medical events, unintended illness or injury, or unfortunate clinical signs in subjects, users, or others related to the investigational product or otherwise.

Discussion

A reliable antenatal age is a prerequisite for preterm newborn classification in birth care settings and constitutes the first step to delivering the necessary care, considering the risks of prematurity.[11] Term newborns, allied with good tone, breathing or crying, are essential elements to determine steps of newborn resuscitation.[21] Although that statement in itself seems very simple, the reality is far from it. Without the certainty of the day in the female cycle on which conception occurred, ultrasound measurement of the crown-rump length is a standard consensual reference for pregnancy dating.[8] This dependence on early echographic scans has deprived many pregnant women and their babies of reliable gestational age.[10] Such a technological gap causes even more disparities than the difference between childbirth scenarios in fully equipped facilities and those ill equipped with scarce technology. Moreover, it can impair the correct classification of infants as premature or growth-restricted.[22,23] The main contribution of this clinical trial is to validate a new approach for gestational age estimation independent of fetal ultrasound measures by demonstrating highly accurate outcomes. Based on two pieces of data—birth weight and ACTFM exposure—and the use of a frugal medical device to assess skin maturity and process algorithms, 359 of 366 preterm neonates with less than 37 weeks of gestation were detected, with 96% being correctly classified.

In this combined study covering enhancing and validating prediction models, we believe the application of k-fold cross-validation with the use of machine learning algorithms provided accurate predictions.[24] While large data samples are unavailable, the process of training and testing are able to estimate the performance of algorithms until we have finished other ongoing clinical trials for external validation.[25] Furthermore, the quantification of uncertainty intervals regarding the predicted gestational age (calculated in days) and comparisons with established references allowed the simulation of real

scenarios for application. Furthermore, the confidence intervals accompanying AUROC's accuracy contributed to revealing the forecast's limits as to discriminating term from preterm newborns at different cutoff points, with clinical relevance. Such strengths are critical to assure the potential value of *the new test* in facing the challenges of postnatal prematurity identification.[26] A new type of data science algorithm has thus emerged with the aim of qualifying pregnancy dating. High-performance reports using learning models based on antenatal ultrasound predictors[27] contradistinguished meager outcomes from those using other morphometric postnatal predictors.[28] Moreover, valuable algorithms with postnatal combinations on maturity scores of newborns are promising, demanding special skills to apply them.[29] Underqualified birth attendants represent a challenge in developing countries, further limiting the use of existing birth care solutions.[30] One advantage of our new device is the skin assessment automation that alerts measurement error caused by movement of the newborn or examiner. Previous reports have detailed the human skin's light-skin interaction and optical properties that benefit this technology.[19,31]

The predictive algorithms used information that health professionals could quickly obtain in childbirth settings—the birth weight and the ACTFM exposure—and which could add value to the physical data of skin maturity. Moreover, the device is capable of providing a gestational age to overcome extreme situations without antenatal records to obtain ACTFM exposure information. However, the algorithm of Model 1 had more comprehensive Bland-Altman 95% limits of 28.6 days compared to the 19.6 days of Model 2 with the full three variables. Working with more flexible forecasts within seven days of error range, 98.7% of newborns had a valuable prediction with Model 2. Considering the simulated scenario with absent or unreliable LMP and lack of ACTFM exposure data—57.7% (n=451) of the sample—the algorithm of Model 1 for gestational age estimation detected 180 of 199 (90.5%) preterm newborns, even with low specificity (48.0%). This result expands the context of use for this medical device since the gestational age based on memory recall of the LMP missed 69 out of 199 preterm newborns, expressing a lower sensitivity (65.8%) when we applied the intent-to-discriminate analysis.

The choice and analysis of two algorithms for gestational age prediction (Model 1 and Model 2) depended on uncertainties regarding the effect of corticosteroids on forecasts. Antenatal corticosteroids to improve newborn outcomes are a practical, evidence-based intervention recommended for women at risk of preterm birth.[32] However, even with the acceleration of lung maturity, the effect of the drug occurs in other organs. The early fetal presence of receptors of corticosteroid hormone receptors in skin epithelial cells indicates that glucocorticoids may play an important role in the differentiation and development of human skin.[33] However, clinical evidence of the effect of ACTFM exposure on skin maturity is weak, and the topic remains unsubstantiated.[34] Thus, until proven otherwise, we interpreted that the importance of ACTFM exposure data to better adjust the gestational age modeling is related to an effect on skin maturity. Even so, we cannot deny that antenatal exposure to corticotherapy is more common in premature infants—72.3% of preterm newborns in this sample. In this respect, this predictor variable could imply a bias favoring preterm newborn detection. The aforementioned ongoing study for external validation of the algorithms could further elucidate this issue because the enrollment process of

newborns introduced the Mozambican birth scenario, where unfortunately ACTFM is not guaranteed for every woman at risk of preterm birth.[25]

Birth weight is a known estimator of risks to newborns. As part of primary routines in childbirth settings, this information has practical applicability, even in facilities with scarce high-cost technologies.[1] Meanwhile, predicting preterm birth based on birth weight when lacking a gold standard is an imperfect solution.[9] Additionally, the LMP reference and the postnatal scores of newborn maturity have demonstrated low accuracy in determining gestational age and identifying prematurity.[35] Later prenatal care and unqualified date recollection justify efforts to enhance the reliability of pregnancy dating through more accurate and accessible technologies, seeking to improve pregnancy outcomes and neonatal survival.[10] In our study, qualifying the LMP at birth with questions about memory of date, menstrual cycles, and checking antenatal clinical documents at birth provided a gestational age able to identify 160/167 (95.8%) preterm newborns. Current approaches to calculating gestational age are sensitive to data quality, resulting in misplaced classification of prematurity.[9] The present study was committed to representing a real scenario in terms of data quality, as stated in the research protocol, with data collection and curation to assure the best reference and comparators for the analysis. Before opening the blinding of the trial, a consistency process confronted data entries with digital images of clinical documents taken during the enrollment. Furthermore, dedicated software was developed exclusively for the clinical trial, considering the quality of the variables and their constraints. Part of the enrollment occurred during the COVID-19 pandemic, causing a minimal amount of missing data, such as ACMF information (4/781 newborns).

Regarding the generalizability of outcomes, this multicenter trial gathered perinatal centers from the northern, central, southwestern, and southern regions of Brazil. This collaborative evaluation contributed to sampling a mixed population of newborns with high miscegenation and involved 15 examiners who attended good clinical practice training. The intraobserver error and interobserver error of measures were slow, corroborating previous results.[19] The number of preterm newborns was enough to analyze subcategories of prematurity as extreme preterm (n=42); however, the overall rate of preterm newborns was 46.9%, values observed in referral facilities and not in the general population of Brazilian newborns.[36] The number of neonatal deaths during 72 h of follow-up was 14, with 12 deaths occurring in newborns with gestational age below 28 weeks due to extreme prematurity complications. We expect to target worse childbirth scenarios for this technology implementation.[30] In addition, the safety of this device is similar to other optical technologies already used in neonatal care, such as transcutaneous bilirubinometer and pulse oximetry.[21,37]

With regard to limitations, the adoption of the new test deserves caution. Impaired fetal growth and large-for-gestational age newborn influence were not included in the analysis. Nevertheless, photometer-based technology for skin maturity assessment has a basis on skin transparency in part associated with epidermal thickness.[19] In a previous analysis of high-frequency ultrasound of newborn skin, we reported that the epidermal layer had no significant influence on the fetal growth pattern when associated with

gestational age.[38] Furthermore, the epidermal thickness of newborns was itself a predictor of gestational age when analyzing postmortem histological skin slices from the sole.[39] For the future, comparisons are expected based on postnatal approaches for gestational age estimation, such as scores of maturity and foot length or image combinations.[28]

Identifying preterm newborns is the first step to attending to their needs. The global rate of neonatal mortality corresponds to 6,700 neonatal daily deaths, mostly from preventable or treatable conditions in scenarios of healthcare scarcity.[40] We hope that strengthening the data sources of healthcare facilities with a reliable gestational age can help in identifying vulnerable newborns in situations with the absence or lack of such information.

Methods

Study design and Participants

A multicenter prospective clinical trial intention-to-diagnosis study by singlegroup, singleblinding, and singlearm with a reference standard. This article adheres to the Transparent Reporting of a Multivariable Prediction Model for Individual Prediction or Diagnosis (TRIPOD) for completeness and clearness.[41] To assess the risk of bias and applicability, the development and validation methods followed guidance from the Prediction Model Risk of Bias Assessment Tool (PROBAST).[42] The clinical trial protocol is disclosed in the Brazilian Clinical Trials Registry (ReBec) RBR-3f5bm5.

This report examined primary and secondary outcomes of data concerning prematurity prediction and safety. Other secondary outcomes related to lung maturity prediction are under analysis for further publication. Five Brazilian referral centers participated in the study: Clinical Hospital – Universidade Federal de Minas Gerais (as coordinator); Hospital Sofia Feldman – Minas Gerais State; Hospital da Universidade Luterana do Brasil – Rio Grande do Sul State; Hospital Materno-infantil de Brasília – Distrito Federal; and Hospital Universitário da Universidade Federal do Maranhão – Maranhão State. The local independent ethics review board of each approved the research protocol, registered under number CAAE 81347817.6.1001.5149 at the Brazilian National Research Council. In addition, parents signed an informed consent form on behalf of the newborn before participating.

A prospective concurrent and sequential process enrolled newborns during the first 24 hours of life. The first enrollment occurred in 2019-01-02, and the last enrollment occurred in 2021-05-30. Eligibility criteria, participants' timeline, and the procedures followed the clinical protocol.[20] In short, we assessed the skin maturity of live newborns with at least 24 weeks or more of gestational age. All of them had their gestational age estimated by clinical parameters, with the embryo measurement mandatorily assessed by ultrasound exam at <14 weeks of gestation. Anhydramnios, hydrops, congenital skin diseases, or chorioamnionitis were the exclusion criteria due to their potential to modify the skin structure.

Procedures

The coordinator center trained 15 examiners following good clinical practice as set forth by the Brazilian regulatory health agency's recommendations.[43] In loco reliability of examiners was assessed during the certification visit of a senior researcher in the collaborator centers (Supplementary Note 3). Standard operating procedures were mandatory to advise enrollment. Data collection was conducted with the use of a smartphone and skin assessment with the medical device under validation.[43,45]

An automated algorithm in the data collection system,[46] blinded to the examiner, calculated the R-gestational age. For this, a set of rules adjusted the due date using data entries collected from the ultrasound report or prenatal care book/document, providing the best reference.[8] Clinical information was collected in structured questionnaires using a software program dedicated to this project. The framework of clinical variables and skin acquisitions is available in Supplementary Note 4. For data curation, the investigator's data entries were confronted with information from photographed digital images of clinical documents. In the case of multiple birth gestations with different ultrasonographic crown-rump-length values, the average of each embryo/fetal value was considered. A double-check system—paper-based and electronic—allowed the verification of the reliability and validity of clinical data as well as skin thickness acquisition. In addition, the data quality of antenatal pregnancy dating was evaluated by comparing the frequency of days in dates, as they should be random with no preference for digits. For this, in cases of multiple gestations, we retained only the first twin information for day-digit evaluation (Supplementary Note 1).

The Intervention

The intervention is a new device that processes the backscattered signal acquired from the skin of the newborn's sole with clinical variables to predict gestational age. Its development included steps from the workbench to the clinical experimentation already described.[19] Likewise, we previously analyzed the best body position to assess the skin reflectance for pregnancy dating, as well as environmental influences such as humidity, temperature, ambient light, and the newborn's skin color.[19,31,38] Regarding the characteristics of the components, wavelengths from 400 nm to 1200 nm of the light emitter categorized the safety level of this medical device as Class II (noninvasive and medium risk) according to the regulatory agency in Brazil.[44] When the light emitting sensor touches the skin over the sole for a few seconds, it triggers 10 automated measurements. The device emits alerts regarding measurement errors caused by involuntary movement of the newborn or examiner under a set of known constraints of the skin reflectance of a newborn.[19] Such an event requires a new attempt. The device output was blinded to the examiners.

The skin assessment occurred with the newborn inside incubators, incubators-radiant warmer, warming pad-bassinet, standard crib, or in the mother's lap to ensure minimum manipulation and avoid unbalancing the clinical conditions. The sensor touched the sole three times, following complete disinfection with alcohol. Fourteen minimum viable products were produced for the study (Fig. 6). At the beginning and the end of the clinical trial, the irradiance emitted by each device and the reflection against

a standard-white Spectralon® offered values for calibration. The adjusted value was the raw value of the acquisition divided by the standard-white reflection.

Improvements on the gestational age estimation model

The new test has an algorithm to predict gestational age, previously described elsewhere and patented.[19,47] However, the current data set groundwork for improvements in the algorithm for predicting gestational age based on the ridge linear regression method[48] to optimize the latent collinearity and avert overfitting variables. Additionally, the nonlinear machine learning method eXtreme Gradient Boosting (XGBoost)[49] created at most 50 trees with a maximum depth of three. Both models were validated using a 10-fold cross-validation approach ten-fold, repeated 30 times. Missing data had no imputation. The analytical pipeline is detailed in Supplementary Note 5.

All clinical variables used as predictors of the models were available at the test time. Therefore, they could be used in real scenarios from user input into the medical device interface. The standalone newborn skin reflectance value was the raw variable, adjusted by clinical variables. Skin reflectance had a strong positive correlation with R-gestational age; Pearson coefficient = 0.79, $p < 0.001$, with a mean absolute error (MAE) of 2.0 weeks (Supplementary Table in Note 2). However, skin acquisition adjusted for birth weight and antenatal use of ACTFM achieved the best performance in gestational age prediction. The set model performance is presented as a Supplementary Table in Note 2, including intermediary analysis with incubator stay, gender, and jaundice variables to verify the elimination of intervenient variables after technological improvements since the early version of the device.[19, 47]

Outcomes

The primary outcome was the agreement between the gestational age predicted by the new test and the R-gestational age in terms of the most accurate predicted models.

A secondary endpoint was the accuracy of the new test to discriminate preterm newborns considering thresholds at 37, 32, and 28 weeks of pregnancy. Moreover, the proportion of preterm newborns correctly detected at birth within a one-week error margin. Another secondary endpoint was comparing the difference between predicted gestational age and the gestational age calculated by a second ultrasound exam after 13 weeks and 6 days of gestation and before 22 weeks via Comparator 1 (C1) and with the LMP, Comparator 2 (C2). This outcome was intended to simulate the performance of the test in scenarios without the reference and compare the agreement between established methods for gestational age calculation and the new test.

The safety of the new device is still a derived endpoint, which refers to unexpected medical events, unintended illness or injury, or unfortunate clinical signs in subjects, users, or others, whether or not they are related to the investigated product.

Statistical analysis

A sample of 787 newborns was required to determine an effect size of 10% to test the hypothesis of equivalence between gestational ages estimated by test and reference, as detailed in our research protocol.[43] Intention-to-diagnosis analysis guaranteed all newborns who were included in the statistical analysis, regardless of any result of the new test, and even inconclusive ones. Descriptive analysis of the newborn's clinical characteristics and the intervention measurements were performed.

Regarding the primary endpoint, the agreement among different methods for gestational age (R, new test, C1, C2) was calculated using the intraclass coefficient correlation and Bland & Altman plots.[50] Regarding the accuracy of predicted gestational age by the models in identifying premature newborns, the area under the receiver operating characteristic curve (AUROC) at 95% CI described the model's discrimination. The significance level for hypothesis tests is 5%, together with 95% confidence intervals.

Declarations

Funding

This research was supported by the Grand Challenges Exploration from the Bill & Melinda Gates Foundation (Grant number OPP1128907, Contract), http://www.gatesfoundation.org/, and *Fundação de Amparo a Pesquisa de Minas Gerais*, Brazil, http://www.fapemig.br/en/, nonprofit sectors. The clinical trial is funded by the Brazilian Ministry of Health, Program of Development of the Industrial Health Complex (PROCIS), project 23072.052747/2017-51. ZSNR is a researcher with a grant from *Conselho Nacional de Pesquisa* (CNPq), Brazil. The funders played no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

The authors would like to thank the Center for Clinical Research of Hospital das Clínicas, Universidade Federal de Minas Gerais for their general support and good clinical practice training and Fotini Toscas from the Brazilian Ministry of Health for active intermediation as trial sponsor contact. We also thank all the clinical staff at the involved hospitals who cared for the patients and undergraduate students who helped with the data collection.

Competing interests

The authors declare a patent deposit on behalf of the Universidade Federal de Minas Gerais and Fundação de Amparo a Pesquisa de Minas Gerais, Brazil, http://www.fapemig.br/en/. The inventors were ZSNR, RNG, BR1020170235688 (CTIT-PN862). BirthTech, a spin-off company, received a license to produce and commercialize this technology, and RNG is its founder.

Data Availability Section, Responsibility, and Analysis

The lead authors (ZSNR, RMCR, RNG, JSG and RAPLA) had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Correspondence and requests for materials should be addressed to ZSNR. Data are available upon reasonable request and after anonymization to allow sharing of data ethically and legally, preserving the confidentiality of the persons who participated in this study.

Word Count: 2876 words

Supplementary Information:

Supplementary Note 1: Gestational age data quality evaluation.

Supplementary Note 2: Correlation between reference gestational age and predictor variables.

Supplementary Fig. 1: Receiver operating characteristic of skin reflection to discriminate term from preterm newborns

Supplementary Note 3: The reliability of the skin assessment with the photometer of the device

Supplementary Note 4: Framework of clinical variables collected from each newborn

Supplementary Note 5: Analytical pipeline

References

1. Tunçalp, Ö. *et al.* WHO recommendations on antenatal care for a positive pregnancy experience—going beyond survival. *BJOG Int. J. Obstet. Gynaecol.* **124**, 860–862 (2017).

2. Lawn, J. E. *et al.* Every Newborn: progress, priorities, and potential beyond survival. *Lancet Lond. Engl.* **384**, 189–205 (2014).

3. Rani, M., Bonu, S. & Harvey, S. Differentials in the quality of antenatal care in India. *Int. J. Qual. Health Care J. Int. Soc. Qual. Health Care* **20**, 62–71 (2008).

4. Walani, S. R. Global burden of preterm birth. *Int. J. Gynecol. Obstet.* **150**, 31–33 (2020).

5. Oza, S., Cousens, S. N. & Lawn, J. E. Estimation of daily risk of neonatal death, including the day of birth, in 186 countries in 2013: a vital-registration and modelling-based study. *Lancet Glob. Health* **2**,

e635-644 (2014).

6. Howson, C. P., Kinney, M. V., McDougall, L., Lawn, J. E., & the Born Too Soon Preterm Birth Action Group. Born Too Soon: Preterm birth matters. *Reprod. Health* **10**, S1 (2013).

7. Kim, E. T., Singh, K., Moran, A., Armbruster, D. & Kozuki, N. Obstetric ultrasound use in low and middle-income countries: a narrative review. *Reprod. Health* **15**, 129 (2018).

8. ACOG. Committee Opinion No 700: Methods for Estimating the Due Date. *Obstet. Gynecol.* **129**, e150–e154 (2017).

9. Miller, L. *et al.* Working with what you have: How the East Africa Preterm Birth Initiative used gestational age data from facility maternity registers. *PLOS ONE* **15**, e0237656 (2020).

10. Karl, S. *et al.* Preterm or not–an evaluation of estimates of gestational age in a cohort of women from rural Papua New Guinea. *PloS One* **10**, e0124286 (2015).

11. WHO. *WHO recommendations on interventions to improve preterm birth outcomes*. (World Health Organization, 2015).

12. Kullinger, M., Granfors, M., Kieler, H. & Skalkidou, A. Discrepancy between pregnancy dating methods affects obstetric and neonatal outcomes: a population-based register cohort study. *Sci. Rep.* **8**, 6936 (2018).

13. WHO. *WHO compendium of innovative health technologies for low-resource settings, 2016- 2017*. (2018).

14. Nelson, G. A. & Holschuh, C. Evaluation of Telehealth Use in Prenatal Care for Patient and Provider Satisfaction: A Step Toward Reducing Barriers to Care. *J. Nurse Pract.* **17**, 481–484 (2021).

15. Meskó, B., Drobni, Z., Bényei, É., Gergely, B. & Győrffy, Z. Digital health is a cultural transformation of traditional healthcare. *mHealth* **3**, 38 (2017).

16. Thompson, M. The geographies of digital health – Digital therapeutic landscapes and mobilities. *Health Place* **70**, 102610 (2021).

17. Miller, D. D. & Brown, E. W. Artificial Intelligence in Medical Practice: The Question to the Answer? *Am. J. Med.* **131**, 129–133 (2018).

18. WHO. *Global strategy on digital health 2020-2025*. (World Health Organization, 2020).

19. Reis, Z. S. N., Vitral, G. L. N., de Souza, I. M. F., Rego, M. A. S. & Guimaraes, R. N. Newborn skin reflection: Proof of concept for a new approach for predicting gestational age at birth. A cross-sectional study. *PloS One* **12**, e0184734 (2017).

20. Reis, Z. S. N. *et al.* Prematurity detection evaluating interaction between the skin of the newborn and light: protocol for the preemie-test multicentre clinical trial in Brazilian hospitals to validate a new medical device. *BMJ Open* **9**, e027442 (2019).

21. Wyckoff, M. H. et al. Neonatal life support: 2020 international consensus on cardiopulmonary resuscitation and emergency cardiovascular care science with treatment recommendations. *Circulation*. **142**, n. 16_suppl_1, p. S185-S221, (2020).

22. Scott, K. *et al.* "I can guess the month … but beyond that, I can't tell" an exploratory qualitative study of health care provider perspectives on gestational age estimation in Rajasthan, India. *BMC Pregnancy Childbirth* **20**, 529 (2020).

23. Wylie, B. J. *et al.* Gestational age assessment in malaria pregnancy cohorts: a prospective ultrasound demonstration project in Malawi. *Malar. J.* **12**, 183 (2013).

24. Weiss, S. M. & Indurkhya, N. Rule-based machine learning methods for functional prediction. *J. Artif. Intell. Res.* **3**, 383–403 (1995).

25. Reis, Z. *et al.* Premature or Small for Gestational Age Discrimination: International Multicenter Trial Protocol for Classification of the Low-Birth-Weight Newborn Through the Optical Properties of the Skin. *JMIR Res. Protoc.* **9**, e16477 (2020).

26. Ananth, C. V. & Brandt, J. S. Fetal growth and gestational age prediction by machine learning. *Lancet Digit. Health* **2**, e336–e337 (2020).

27. Fung, R. *et al.* Achieving accurate estimates of fetal gestational age and personalised predictions of fetal growth based on data from an international prospective cohort study: a population-based machine learning study. *Lancet Digit. Health* **2**, e368–e375 (2020).

28. Torres, M. T., Valstar, M. F., Henry, C., Ward, C. & Sharkey, D. Small sample deep learning for newborn gestational age estimation. in 79–86 (IEEE, 2017).

29. Rittenhouse, K. J. *et al.* Improving preterm newborn identification in low-resource settings with machine learning. *PloS One* **14**, e0198919 (2019).

30. Soubeiga, D., Gauvin, L., Hatem, M. A. & Johri, M. Birth Preparedness and Complication Readiness (BPCR) interventions to reduce maternal and neonatal mortality in developing countries: systematic review and meta-analysis. *BMC Pregnancy Childbirth* **14**, 129 (2014).

31. Silva, P. C., Guimarães, R. N., Souza, R. G. & Reis, Z. S. N. A quantitative cross-sectional analysis of the melanin index in the skin of preterm newborns and its association with gestational age at birth. *Skin Res. Technol. Off. J. Int. Soc. Bioeng. Skin ISBS Int. Soc. Digit. Imaging Skin ISDIS Int. Soc. Skin Imaging ISSI* **26**, 356–361 (2020).

32. Emeruwa, U. N., Krenitsky, N. M., & Sheen, J. J. Advances in Management for Preterm Fetuses at Risk of Delivery. *Clinics in Perinatology*, **47**, 685-703 (2020).

33. Condon, J. *et al.* Expression of type 2 11β-hydroxysteroid dehydrogenase and corticosteroid hormone receptors in early human fetal life. *J. Clin. Endocrinol. Metab.* **83**, 4490–4497 (1998).

34. August, D. & Kandasamy, Y. The effects of antenatal glucocorticoid exposure on fetal and neonatal skin maturation. *J. Perinat. Med.* **45**, (2017).

35. Lee, A. C. *et al.* Diagnostic Accuracy of Neonatal Assessment for Gestational Age Determination: A Systematic Review. *Pediatrics* **140**, e20171423 (2017).

36. Leal, M. do C. *et al.* Saúde reprodutiva, materna, neonatal e infantil nos 30 anos do Sistema Único de Saúde (SUS). *Ciênc. Saúde Coletiva* **23**, 1915–1928 (2018).

37. Bhutani, V. K., Gourley, G. R., Adler, S., Kreamer, B., Dalin, C., & Johnson, L. H. Noninvasive measurement of total serum bilirubin in a multiracial predischarge newborn population to assess the risk of severe hyperbilirubinemia. *Pediatrics*, **106**, e17-e17 (2020).

38. Vitral, G. L. N. *et al.* Skin thickness as a potential marker of gestational age at birth despite different fetal growth profiles: A feasibility study. *PLOS ONE* **13**, e0196542 (2018).

39. de Souza, I. M. F., Vitral, G. L. N., Caliari, M. V. & Reis, Z. S. N. Association between the chronology of gestation and the morphometrical skin characteristics at childbirth: a development of predictive model. *BMJ Health Care Inform.* **28**, e100476 (2021).

40. UNICEF. IGME: Levels and Trends in Child Mortality. *UNICEF DATA* https://data.unicef.org/resources/levels-and-trends-in-child-mortality/ (2020).

41. Collins, G. S., Reitsma, J. B., Altman, D. G. & Moons, K. G. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement. *BMC Med.* **13**, 1 (2015).

42. Venema, E. *et al.* Large-scale validation of the prediction model risk of bias assessment Tool (PROBAST) using a short form: high risk of bias models show poorer discrimination. *J. Clin. Epidemiol.* **138**, 32–39 (2021).

43. BRASIL. Resolução - RDC Nº 10, de 03 de Março de 2015. (2015).

44. Emeruwa, Ukachi N.; Krenitsky, Nicole M.; Sheen, Jean-Ju. Advances in Management for Preterm Fetuses at Risk of Delivery. **Clinics in Perinatology**, v. 47, n. 4, p. 685-703, 2020. born and light: protocol for the preemie-test multicentre clinical trial in Brazilian hospitals to validate a new medical device. *BMJ Open* **9**, e027442 (2019).

45. Reis, Z., Vitral, G. L. N., Guimarães, R. N., Aguiar, R. A. P. L. de & Romanelli, R. M. C. The Preemie-Test for the assessment of the newborn skin maturity. (2019) doi:10.17504/protocols.io.7ynhpve.

46. Reis, Z., Gaspar, J. de S., Elias, S. O. & Aguiar, R. A. L. P. D. Algorithm for gestational age assessment at birth. (2020) doi:10.17504/protocols.io.bawbifan.

47. Zilma, S. N. R. & Rodney, N. G. Device and method for determining gestational age. (2018). patent/US20200178847A1/en.

48. Meijer, R. J. & Goeman, J. J. Efficient approximate k-fold and leave-one-out cross-validation for ridge regression. *Biom. J.* **55**, 141–155 (2013).

49. Chen, T. & Guestrin, C. XGBoost: A Scalable Tree Boosting System. in *the 22nd ACM SIGKDD International Conference* 794 (2016). doi:10.1145/2939672.2939785.

50. Altman, D. G. & Bland, J. M. Measurement in Medicine: The Analysis of Method Comparison Studies. *J. R. Stat. Soc. Ser. Stat.* **32**, 307–317 (1983).

51. Popović, Z. B.; Thomas, J. D. Assessing observer variability: a user's guide. *Cardiovascular diagnosis and therapy*. **7,** 317-324, (2017).

52. Kohavi, R. A study of cross-validation and bootstrap for accuracy estimation and model selection. *Ijcai*. Montreal, Canada. **14**, 1137-1145 (1995).

53. Friedman, J. H. Stochastic gradient boosting. *Computational statistics & data analysis*. **38**, 367-378 (2002).

Tables

Table 1. Baseline characteristics of the pregnancy and newborns

	N	Statistics
Demographic and antenatal data of women		
Maternal age (years), median (IQR)	704	27 (9)
First prenatal care assessment (weeks), median (IQR)	616	9 (4)
Reliable last menstrual period among women with memory of this data, n/N (%)	614	296 (48.2%)
Gestational age at the first ultrasound assessment (weeks), median (IQR)	704	10.1 (3.7)
Gestational age at the second ultrasound assessment (weeks), median (IQR)	704	19.4 (4.3)
Diabetes, n/N (%)	703	104 (14.8%)
Hypertensive disturbance during pregnancy, n/N (%)	704	147(20.9%)
ACMF, n/N (%)	701	215 (30.7%)
Multiple gestation, n/N (%)	704	76 (10.4%)
Newborns at the first day of life		
Gestational age at birth (weeks)	781	37.3 (6.3%)
Preterm[a], n/N (%)	781	366 (46.9%)
Moderate to late preterm[b], n/N (%)	781	235 (30.2%)
Very preterm[c], n/N (%)	781	89 (11.4%)
Extremely preterm[d] , n/N (%)	781	42 (5.4%)
Small for gestational age, n/N (%)	781	116 (14.9%)
Appropriate for gestational age, n/N (%)	781	606 (77.6%)
Large for gestational age, n/N (%)	781	59 (7.6%)
Major malformation, n/N (%)	781	9 (1.2%)
1-minute Apgar score, median (IQR)	775	8 (1%)
5-minute Apgar score, median (IQR)	777	9 (1%)
Sex, male, n/N (%)	781	390 (49.9%)
Birthweight (g), median (IQR)	781	2740 (1498%)
NICU at the skin assessment, n/N (%)=me	781	280 (35.9%)
Jaundice at the skin assessment, n/N (%)	773	249 (31.9%)
Phototherapy at the skin assessment, n/N (%)	751	32 (4.1%)
During 72h hours of life		
Newborn mortality, n/N (%)	781	14 (1.8%)

ACMF: Antenatal Corticosteroid Therapy for Fetal Maturation. NICU: neonatal intensive care unit. IQR: interquartile range.

[a]Less than 37 weeks. [b]More than 32 to less than 37 weeks. [c]More than 28 to less than 32 weeks. [d]Less than 28 weeks.

Table 2. Agreement between predicted gestational age by the models with the established references

	Model 1-gestational age	Model 2-gestational age
Bland-Altman 95% limits (days)	(-28.6 to 28.6)	(-19.6 to 19.6)
ICC with R-gestational age (95% CI)	0.869 (0.851 to 0.885)	0.941 (0.933 to 0.949)
ICC with Comparator 1 (95% CI)	0.873 (0.855 to 0.889)	0.932 (0.922 to 0.941)
ICC with Comparator 2 (95% CI)	0.798 (0.870 to 0.904)	0.860 (0.840 to 0.979)

CI: confidence interval. ICC: intraclass correlation coefficient. R: reference standard

Comparator 1 is the gestational age calculated with a second antenatal ultrasound exam after 13 weeks and 6 days of gestation and before 22 weeks.

Comparator 2 is the gestational age calculated with the last menstrual period.

Model 1: Linear-Ridge skin reflectance and birth weight model predictors for gestational age.

lel 2: XGBoost skin reflectance, birth weight, and ACTFM exposure predictors for gestational age.

Table 2: Descriptive analysis of frequency for gestational age at birth, according to the methods of calculation.

	Median (IQR) days	Min - Max, days
Reference gestational age (n=781)	261 (44)	155 to 294
Comparator 1 gestational age (n=781)	260 (43)	167 to 322
Comparator 2 gestational age (n=680)	262 (48)	165 to 319
Model 1 gestational age (n=781)	258.9 (43)	188.3 to 324.6
Model 1 gestational age (n=777)	259.4 (43.6)	168.6 to 287.7
Difference between Comparator 1 and reference gestational age (n=781)	0 (4)	-27 to 50
Difference between Comparator 2 and reference gestational age (n=680)	-0.01 (4)	-72 to 72
Difference between Model 1 and reference gestational age (n=781)	0.21 (14)	-48.6 to 48.6
Difference between Model 2 and reference gestational age (n=777)	-1.4 (-2.1)	-9.7 to 16.2

IQR: interquartile range.

Comparator 1 is the gestational age calculated with a second ultrasound exam after 13 weeks and 6 days of gestation and before 22 weeks.

Comparator 2 is the gestational age calculated with the last menstrual period.

Model 1: Linear-Ridge skin reflectance and birth weight predictors.

Model 2: XGBOOST skin reflectance, birth weight, and ACTFM exposure predictors.

Table 3: Prediction accuracy for preterm newborn by the methods of gestational age estimation

	Comparator 1 gestational age (n=781)	Comparator 2 gestational age (n=680)	Model 1 gestational age (n=781)	Model 2 gestational age (n=777)
Preterm <37 weeks (n=366)				
ACU (%)	96.2	92.8	87.3	96.0
False - (%)	3.8	9.4	10.1	1.9
SPE (%)	96.1	94.7	85.1	94.2
VPP (%)	95.7	93.9	84.1	93.7
VPN (%)	96.6	91.9	90.5	98.2
LR+	25.0	17.1	6.0	16.8
LR-	0.04	0.10	0.10	0.02
Preterm <32 weeks (n=131)				
ACU (%)	98.5	96.2	91.3	95.0
False - (%)	3.8	10.0	23.7	11.5
SPE (%)	98.9	97.4	94.3	96.3
VPP (%)	94.7	86.8	73.0	82.9
VPN (%)	99.2	98.1	95.2	97.6
LR+	89.3	34.2	13.4	23.8
LR-	0.04	0.10	0.25	0.12
Preterm <28 weeks (n=42)				
ACU (%)	99.0	98.5	95.4	97.2
False - (%)	11.9	16.7	71.4	19.0
SPE (%)	99.6	99.4	99.2	98.1
VPP (%)	92.5	88.2	66.7	70.8
VPN (%)	99.3	99.1	96.1	98.9
LR+	217	134	35.2	42.5
LR-	0.12	0.17	0.72	0.19

Model 1: Linear-Ridge skin reflectance and birth weight predictors. Model 2: XGBOOST skin reflectance, birth weight, and ACTFM exposure predictors. ACU: newborn correctly classified (accuracy). LR+: Likelihood ratio positive. Likelihood ratio negative: LR-. SEN: sensibility. SPE: Specificity.

Table 5. Intent to preterm discrimination according to simulated scenario of care

ıario of care	Scenario-one: Absent or unreliable LMP				Scenario-two: Reliable LMP			
	Preterm newborns	Sens	Spec	ACU	Preterm newborns	Sens	Spec	ACU
rence ıational age 81)	199/451	...	...	...	167/330	...	...	...
el 1-gestational (n=781)	220/451	180/199 (90.5%)	212/252 (48.0%)	391/451 (86.9%)	171/330	149/167 (89.2%)	141/163 (86.5%)	290/330
el 2-gestational (n=777)	212/451	198/199* (99.5%)	234/252*	432/451* (95.8%)	171/330	161/167 (96.4%)	153/163 (93.9%)	314/330 (95.2)
parator 1- ıational age 81)	201/451	190/199 (95.5%)	241/252 (95.6%)	431/451 (95.6%)	167/330	162/167 (97.0%)	158/163 (96.9%)	320/330 (97.0%)
parator 2- ıational age 80)	144/451	131/199** (65.8%)	183/252** (72.6%)	314/451** (69.6%)	166/330	160/167 (95.8%)	157/163 (96.3%)	317/330 (96.1%)

ACU: newborn correctly classified (accuracy). LMP: last menstrual period. *Including inconclusive test in the model 2: three newborns were preterm; one newborn was term. **Including absence of comparator 2 gestational age: 45 newborns were preterm, 56 newborns were term.

Comparator 1 is the gestational age calculated with a second antenatal ultrasound exam after 13 weeks and 6 days of gestation and before 22 weeks.

Comparator 2 is the gestational age calculated by the last menstrual period.

Model 1: Linear-Ridge skin reflectance and birth weight predictors for gestational age.

lel 2: XGBoost skin reflectance, birth weight, and ACTFM exposure predictors for gestational age.

Figures

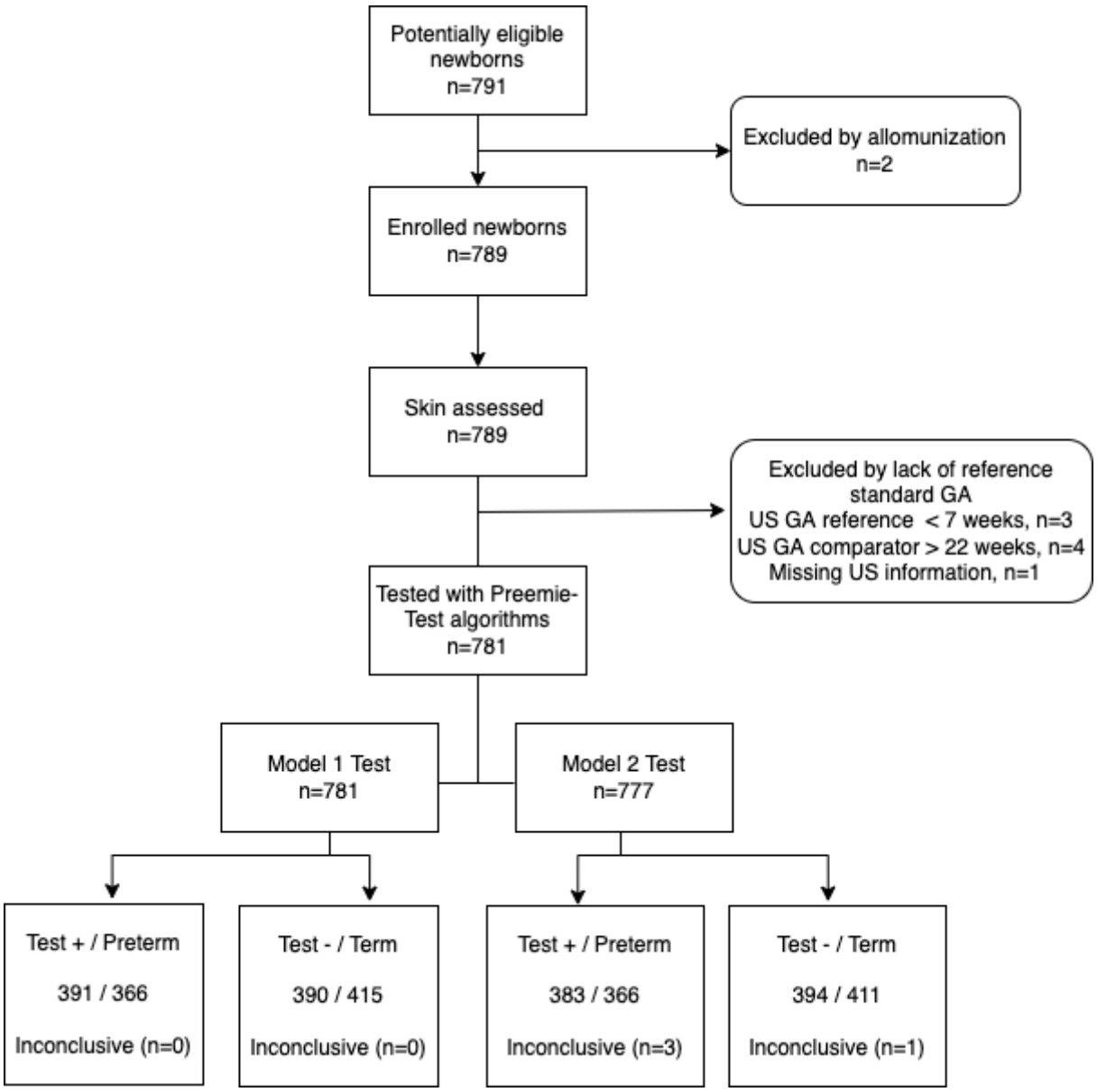

Figure 1

Flow diagram of participants throughout the study with results for the predictive models

Notes: Model 1 Test: Linear-ridge skin reflectance and birth weight predictors for gestational age. Model 2 Test: XGBoost skin reflectance, birth weight, and antenatal corticosteroid therapy for fetal maturation exposure predictors for gestational age.

(+) positive. (-) negative.

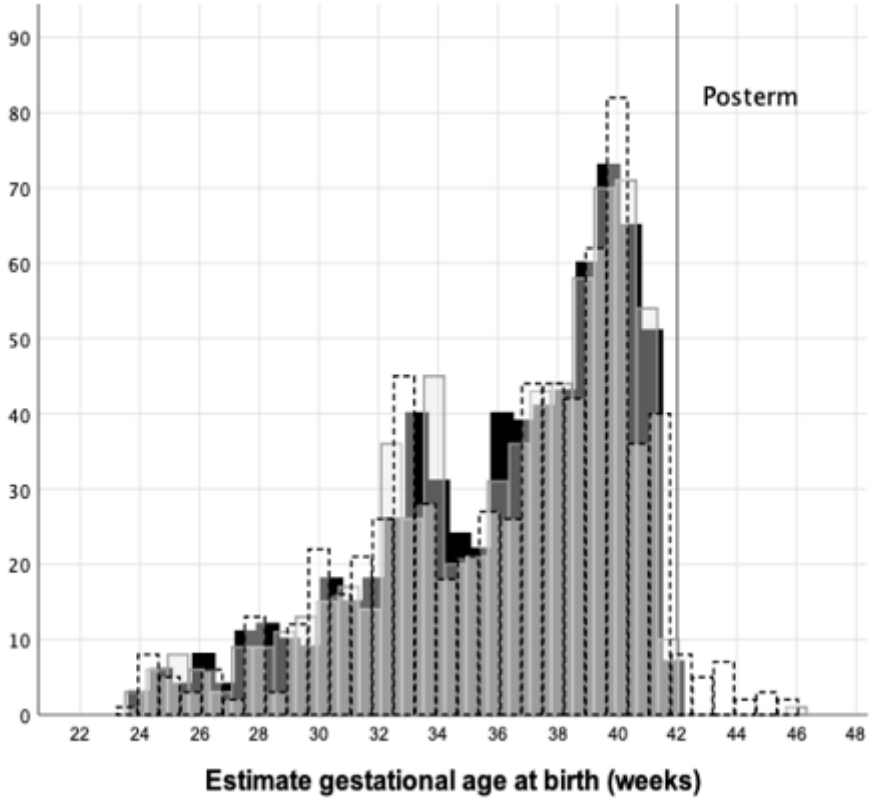

Figure 2

The distribution of estimated gestational age at birth by reference standard (R), ultrasound between 14 and 22 weeks of gestation (C1), and the last menstrual period (C2). Notes: Black boxes: distribution of R-gestational age had a median of 37.3 (IQR 6.3) weeks. Gray boxes: distribution of C1-gestational age had a median of 37.31 (6.1) weeks. Hatched boxes: The distribution of C2-gestational age had a median of 37.4 (6.8) weeks.

Figure 3

Correlation plot between the skin reflectance of the newborn and the reference gestational age at birth. Notes: The hatched lines correspond to 95% of the confidence interval. Black diamonds represent preterm newborns by the reference gestational age. Gray circles: represent term newborns by the reference gestational age.

Figure 4

*When Comparator 2 is available. GA: gestational age

Comparator 1 is the gestational age calculated by a second ultrasound exam after 13 weeks and 6 days of gestation and before 22 weeks.

Comparator 2 is the gestational age calculated by the last menstrual period.

Model 1: Linear-Ridge skin reflectance and birth weight model predictors.

Model 2: XGBoost skin reflectance, birth weight, and ACTFM exposure predictors.

Figure 5

Receiver operating characteristic for the models to discriminate terms from preterm newborns. AUROC: the area under the receiver operating characteristic curve.

Model 1: Linear ridge skin reflectance and birth weight predictors for gestational age, n=781

Model 2: XGBoost skin reflectance, birth weight, ACFM exposure predictors for gestational age, n=777

Figure 6

The new device and its simulated application on a newborn-doll. Source: the authors

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterialUFMG.docx