Published in final edited form as:

Arch Dis Child Fetal Neonatal Ed. 2016 November ; 101(6): F494–F501. doi:10.1136/archdischild-2015-309670.

Gestational Age and Birth Weight for Risk Assessment of Neurodevelopmental Impairment or Death in Extremely Preterm Infants

Ariel A. Salas¹, Waldemar A Carlo¹, Namasivayam Ambalavanan¹, Tracy L Nolen², Barbara J Stoll³, Abhik Das², Rosemary D. Higgins⁴, and the Eunice Kennedy Shriver National Institute of Child Health and Human Development Neonatal Research Network

¹University of Alabama at Birmingham, Birmingham, AL, United States

²RTI International, Research Triangle Park, NC, United States

³Emory University, Atlanta, GA, United States

⁴GDB and FU Subcommittee, NICHD Neonatal Research Network, Bethesda, MD, United States.

Abstract

Background—The risk for poor outcomes in preterm infants is primarily determined by birth weight (BW) and gestational age (GA). It is not known whether BW is a better outcome predictor than GA.

Objective—To test whether BW is better than GA (measured in days, rather than completed weeks) for prediction of neurodevelopmental impairment (NDI) and death.

Design/Methods—Extremely preterm infants born at the National Institute of Child Health and Human Development (NICHD) Neonatal Research Network centers between 1998 and 2009 were

Address correspondence to: Waldemar A. Carlo, MD, Department of Pediatrics, University of Alabama at Birmingham, 1700 6th Ave South, Suite 9380 Women and Infants Center, Birmingham, AL, 35249, wcarlo@peds.uab.edu, 205-934-4680.

Financial Disclosure: The authors have no financial relationships to disclose

Conflicts of Interest: The authors have no conflict of interest to disclose

Contributors' Statements

Ariel A Salas: Dr. Salas conceptualized and designed the study, oversaw data analyses, drafted and revised the manuscript, and approved the final manuscript as submitted.

Waldemar A Carlo: Dr. Carlo conceptualized and designed the study, oversaw data analyses, drafted and revised the manuscript, and approved the final manuscript as submitted.

Namasivayam Ambalavanan: Dr. Ambalavanan assisted in study design, reviewed and revised the manuscript, and approved the final manuscript as submitted.

Tracy L Nolen: Dr. Nolen assisted in study design, performed the statistical analyses, reviewed and revised the manuscript, and approved the final manuscript as submitted.

Barbara J Stoll: Dr. Stoll assisted in study design, reviewed and revised the manuscript, and approved the final manuscript as submitted.

Abhik Das: Dr. Das assisted in study design, performed the statistical analyses, reviewed and revised the manuscript, and approved the final manuscript as submitted.

Rosemary D Higgins: Dr. Higgins assisted in study design, reviewed and revised the manuscript, and approved the final manuscript as submitted.

“The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non-exclusive for government employees) on a worldwide basis to the BMJ and co-owners or contracting owning societies (where published by the BMJ on their behalf), and its Licensees to permit this article (if accepted) to be published in Archives of Disease in Childhood and any other BMJ products and to exploit all subsidiary rights, as set out in our licence.”

studied. For the unadjusted analysis, the associations of GA (in days based on best obstetrical estimate) and BW (in grams) with NDI or death were compared using area under the curve (AUC). Adjusted analyses were performed using birth year, sex, race, antenatal steroids, singleton birth, preeclampsia, Apgar score at 5 minutes, and small for gestational age as covariates.

Results—10,652 preterm infants (89%) had outcome data at 18 to 22 months' corrected age. The mean BW was 678 g (SD: 155) and the mean GA was 173 days (SD: 10) or 24^{5/7} weeks (SD: 1^{3/7}). The AUC for NDI or death was 80% with BW and 79% with GA (p=0.82). Unadjusted and adjusted analyses did not differ. NDI or death rates decreased with increasing GA through 26 weeks (estimated risk reduction with each additional day of gestation: 2.2%)

Conclusions—Both BW in grams and GA in days are good predictors of NDI and death in a preterm population selected on the basis of reliable gestational age.

Keywords

outcome prediction; risk stratification; extremely-low-birth-weight infants; extremely-low-gestational-age newborns; infant, extremely premature; neonatal morbidity; neonatal mortality

Introduction

Gestational age (GA) and birth weight (BW) are the main determinants of poor outcomes in extremely preterm infants¹⁻⁶. As a result, efficacy and outcome data based on either GA or BW are a valuable resource to provide medical care, to counsel parents, and to define eligibility and stratification criteria in clinical trials⁷.

GA is a measure of duration of gestation often considered a surrogate of maturity⁸ that is normally associated with proportional weight gain, and BW is a measure of body mass that normally increases with advancing gestation. This interdependence or collinearity between GA and BW has made both terms interchangeable for risk assessment. However, the selection of one as the best indicator of baseline risk in preterm infants has resulted in conflicting clinical evidence and expert opinion⁸⁻¹⁰.

In order to determine the comparative ability of GA and BW to predict outcomes, factors that affect the interdependence between GA and BW must be considered. Maternal morbidity that induce preterm birth¹⁰ and/or cause fetal growth restriction is one of those factors. Inaccurate estimation of GA is another confounding factor, particularly when availability of ultrasound to supplement clinical and historical dating is limited. The other important consideration for risk assessment with either GA or BW is the scale difference between GA in completed weeks of gestation (ordinal variable) and BW in grams (more nearly continuous variable)¹¹.

The aim of this study was to examine the unadjusted and adjusted association of GA in days and BW in grams with neurodevelopmental impairment (NDI) at 18 to 22 months' corrected age or death in extremely preterm infants. We hypothesized that BW would be a better predictor of long-term neonatal outcomes than GA, even after conversion of GA into a more nearly continuous variable.

Methods

Preterm infants born at one of the National Institute of Child Health and Human Development (NICHD) Neonatal Research Network (NRN) centers who were admitted for neonatal care or died in the delivery room prior to admission and met criteria for the NRN registry and follow-up studies between 1998 and 2009 were included. Data from all NRN centers that participated during the study period were extracted. Extremely preterm infants are routinely followed up to 18 to 22 months' corrected age in these centers, but entry criteria for follow-up changed over time. Before 2007, infants were eligible for follow-up if their BW was less than 1000 g. Since 2007, only infants with GA less than or equal to 26 completed weeks of gestation or those enrolled in a randomized trial or approved observational study in the preterm population are followed. For this study, infants were excluded if they were more than 26 completed weeks of gestation, had a major anomaly, or were born to mothers without prenatal care (the best obstetrical estimate is less accurate if ultrasound measurement of fetal anthropometry is done closer to delivery).

The primary composite outcome of this study was NDI at 18 to 22 months' corrected age among survivors or death at any time during hospitalization or before the 18 to 22 months' corrected age follow-up visit. Individual components of the primary composite outcome were considered secondary outcomes. Other secondary outcomes included bronchopulmonary dysplasia (BPD) defined as oxygen use at 36 weeks, severe intraventricular hemorrhage (IVH, grade 3 or 4), periventricular leukomalacia (PVL), and their respective combinations with the outcome death.

For infants born up to 2005, NDI at 18 to 22 months' corrected age was defined as 1 or more of the following: a Bayley Scales of Infant Development (BSID) II Mental Developmental Index (MDI) score of less than 70 (2 SDs below the mean for normal infants), a BSID II Psychomotor Developmental Index (PDI) score of less than 70, moderate to severe cerebral palsy with gross motor function level of 2 or greater, blindness (no useful vision in either eye), or deafness (functional hearing impairment with aids in both ears)¹².

For infants born after 2005, NDI was defined as 1 or more of the following: BSID III cognitive composite score of less than 85, gross motor function level of 2 or greater, blindness, or deafness (functional hearing impairment)¹³.

GA was determined according to standard NRN definitions using the best obstetrical estimate. SGA was defined as BW below the 10th percentile for GA in weeks according to recent growth curves¹⁴. For regression modeling purposes, estimated GA in weeks/days was converted to a more nearly continuous scale (days).

Statistical analyses

All infants were assessed for eligibility. The Spearman correlation coefficient between GA and BW was estimated to investigate collinearity. Missing values were handled with sensitivity analyses using multiple imputation techniques (SAS PROC MI procedure for missing outcome data conditional on subject demographic and baseline covariates and MCMC for covariates in which the monotone missing assumption for outcome data was

held). For internal validation, cross-validation was developed to estimate accurately the performance of the final predictive model. Specifically, infants eligible for analyses were randomly divided into training (50%) and validation (50%) datasets. Initial analyses and modeling decisions were made using the training dataset with the final analyses repeated on the validation dataset to assess consistency.

For the unadjusted analysis, GA and BW were introduced individually in a nominal logistic regression model with NDI or death as the dependent variable. The predictive power of the relationship between GA (in days) and NDI or death and the relationship between BW and NDI or death were compared using the area under the curve (AUC)^{7, 11} with the assumption that a greater AUC would indicate a better discriminating ability of the model (0.60 – 0.70 = poor; 0.71 – 0.80 = fair; 0.80 to 0.90 – good). A similar procedure was conducted for secondary outcomes.

For the adjusted analysis, the following covariates were included: birth year, sex, race (using 3 categories: black, white, and other), exposure to antenatal corticosteroids, singleton birth, preeclampsia, Apgar score at 5 minutes, and small for gestational age (SGA, < 10th percentile)^{5, 7}. These covariates were included in the analysis to adjust for factors that could influence fetal growth¹⁵ and outcomes. For the adjusted analysis, GA and BW were introduced separately in two different models (models A and B) in order to determine the statistical significance and overall performance of GA and BW in the presence of other covariates⁹. In model A, NDI or death was the dependent variable and BW plus covariates were the independent variables. In model B, NDI or death was the dependent variable and GA in days plus covariates were the independent variables. The AUCs obtained from the two models were compared. The individual components of the primary composite outcome and the other individual and composite secondary outcomes were analyzed using a similar approach. In addition, a pre-specified subgroup analysis including only singleton infants was performed based on reports suggesting up to 10% greater risk of NDI or death in twin infants^{16, 17}.

Models with polynomial terms for GA, BW, and birthweight z-scores were explored to quantify the effect of outliers in the primary linear regression model and to determine whether these alternative models would provide a better prediction of NDI or death.

All data were analyzed by the NRN Data Coordinating Center at RTI International using SAS 9.2 (Cary, NC).

Results

The study population included 12,025 infants born in 19 NRN centers. Among these infants, 89% had known outcome of NDI or death and were included in the analysis (n=10,652). Baseline characteristics of the study population are shown in **Table 1**. The mean BW was 678 g (SD: 155) and the mean GA was 173±10 days (24^{5/7} SD ± 1^{3/7} wks). Training and validation data sets (n=5329 and 5323 respectively) had comparable baseline characteristics.

Histograms of the raw relationship between BW and NDI/death and between GA and NDI/death were created using the training dataset (**Figure 1**). For these histograms, the rate of

NDI or death was calculated for subjects within 11 intervals of BW and 11 intervals of GA. The rate of NDI or death decreased as BW and GA increased. The overall correlation coefficient between BW and GA was 0.67 (Spearman, $p = <0.0001$) (**Figure 2**).

The AUC comparison of unadjusted and adjusted analyses for the primary outcome of NDI or death in training and validation sets is shown in **Table 2**. The unadjusted AUC for prediction of NDI or death did not differ between BW and GA ($p=0.48$). In the adjusted analyses, AUC for prediction of NDI or death was 80% with BW in grams and 79% with GA in days ($p=0.82$). The introduction of BW, GA, or both within the adjusted model consistently increased the ability to predict the outcome NDI or death. Additionally, AUC values of adjusted models with both BW and GA were similar to AUC values of unadjusted models with either BW or GA alone (**Table 2**). A sensitivity analysis with multiple imputation techniques for binomial data employed to impute values for missing NDI or death data conditional on demographic and baseline covariates (those included in the adjusted analyses)¹⁸ also showed that the AUC for prediction of NDI or death at 18 to 22 months' corrected age did not differ between BW and GA (unadjusted $p=0.47$; adjusted $p=0.82$). The adjusted analyses were also subject to sensitivity analyses with exclusion of SGA as this covariate relies on both BW and GA. Exclusion of SGA as covariate in the adjusted analysis resulted in a non-significant difference in the ability of GA over BW to predict NDI or death (80% vs 78%, respectively; $p=0.19$).

The possibility of a non-linear association between GA or BW and NDI or death was explored with polynomial regression models. The AUCs resulting from these complex models were not significantly different from the AUCs obtained in linear models and did not provide a better fit (**supplementary material, eTable 1**). Similarly, when birthweight z-scores were used instead of BW controlling for GA¹¹, the resulting AUCs were not significantly different to those obtained in regression models with BW.

The subgroup analysis including only singleton infants generated similar results to those obtained in the overall cohort (**supplementary material, eTable 2**).

The AUC comparison of unadjusted and adjusted analyses for individual components of the primary outcome and secondary outcomes are shown in eTable 3 (supplementary material) and **Table 3**, respectively. The unadjusted and adjusted AUCs for prediction of NDI with both BW and GA were poor and comparable between BW and GA ($p=0.35$). For the outcome NDI, neither BW nor GA added predictive value to the other immediate postnatal variables. Death through follow-up as well as IVH and the composites of IVH and PVL with death were better predicted with GA in unadjusted analyses (AUC unadjusted p values < 0.05); however, superiority of GA over BW on prediction of these outcomes was not observed in the adjusted analyses. In contrast, the unadjusted and adjusted AUC for BW as a predictor of BPD were slightly higher than the unadjusted AUC (63% vs 60%; $p<0.01$) and adjusted AUC for GA (66% vs 63%; $p=0.02$).

Discussion

This study compared the ability of BW and GA to predict NDI or death in extremely preterm infants. Contrary to our hypothesis, BW alone was not a better predictor than GA alone. AUCs for BW and GA as predictors of NDI or death did not differ in unadjusted and adjusted analyses. The inclusion of other variables added very little to the overall performance of the unadjusted model. BW performed better than GA for prediction of the outcome BPD. GA in days and the outcome NDI or death had a linear relationship within the GA range relevant to extreme prematurity. This finding suggests an evenly distributed 2.2% risk reduction of NDI or death with each additional day of gestation.

One of the limitations of this study was the algorithm used to estimate GA. In general, neonatal estimates of GA are less precise than obstetric estimates⁹, particularly in infants born at 28 weeks of gestation or less¹⁹. The predefined NRN preference to record the best obstetrical estimate over the neonatal estimate and the inclusion of only infants born to mothers with prenatal care helped us to ensure that GA was estimated predominantly by an accurate obstetric assessment and not simply by neonatal assessment. Examining GA in a more nearly continuous scale with smaller increments in time of gestation (i.e. days)²⁰ also reduced this limitation and strengthened the overall precision of the analysis.

The other two important limitations of the study were the changes in the eligibility for NRN follow-up during the study period (i.e. overrepresentation of SGA infants before 2007), and the changes in the definition of NDI that occurred during the follow-up period. Because BSID III cognitive scores are approximately 7 points higher than BSID II MDI scores^{21, 22}, higher cutoff scores are recommended to define severe cognitive impairment with BSID III²². We used the alternative definition of NDI (BSID III cognitive score < 85) to overcome this limitation.

One of the strengths of our analysis was the assessment of non-linearity with polynomial regression models²³. Predictive variables such as GA and BW examined in continuous scales tend to imply linearity with outcomes such as mortality⁸. Although most variables do not have a strictly linear relationship with outcomes^{4, 23}, our regression models showed that the relationship between GA and NDI or death throughout the entire GA range relevant to extreme prematurity is linear. This finding suggests that the benefit of a 1-day increment in GA for risk reduction of NDI or death is constant throughout extreme prematurity and not restricted to GA at the limits of viability²⁰.

Although the inclusion of liveborn infants born at centers that offer universal resuscitation and centers that offer selective resuscitation at 22 or 23 weeks gestation increases the external validity of our results, it was beyond the scope of our study to determine whether knowing GA has a greater effect than knowing BW on decisions to provide care for liveborn infants.

Previous studies have shown that selective exclusion of both BW and GA from regression and neural network models decreased significantly the AUC for prediction of mortality⁵. It has also been reported that each additional week of GA increases survival to discharge without major morbidities⁶, decreases the risk of NDI and death⁷, and remains a valuable

parameter for prenatal counseling^{6, 24}. Nevertheless, prediction of NDI or death with GA in weeks alone is less accurate than predictive models with GA, sex, exposure to antenatal corticosteroids, singleton birth and BW^{7, 25}. This difference in accuracy and precision between models suggests that mortality risk attributable to immaturity alone is moderate (roughly 50% in simulated epidemiologic studies)¹⁰ and that BW is important in the assessment of neonatal risk²⁶ as it represents a proxy measure of the interaction between immaturity (i.e. GA) and other processes that cause preterm delivery (i.e. maternal morbidity)¹⁰.

It is certainly plausible that describing high-risk populations in samples defined by BW alone may have induced an over-representation of growth retarded/restricted infants^{6, 8, 27} and masked the risk inference attributed to immaturity by amplification of some causal factors²⁴; however, BW is subject to less measurement bias than GA. Measurement errors can occur when GA is based on maternal recall of the last menstrual period²⁷ and also after early ultrasound examinations^{9, 23, 27, 28} which can account for up to 1 to 2 weeks of difference in many extremely preterm infants^{9, 23}.

In summary, our results indicate that BW in grams and GA in days are both good predictors of NDI and death in a preterm population with reliable gestational ages. Our results also show an evenly distributed risk reduction of NDI or death with each additional day of gestation from 23 through 26 weeks gestation. Although it is possible that BW matters more than maturity for some outcomes such as BPD⁴, our results indicate that both are good outcome predictors. Prognostic stratification of baseline risk for NDI or death in extremely preterm infants prior to enrollment on neonatal studies can be achieved based on either BW or GA. Results from research conducted with enrollment criteria primarily based on BW < 1000 grams can only be extrapolated to results from research conducted in preterm infants 27 weeks of gestation when neonatal estimates of GA or other alternative methods for gestational dating can confirm that infants with GA > 27 weeks of gestation have been excluded.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgment

The National Institutes of Health and the Eunice Kennedy Shriver National Institute of Child Health and Human Development provided grant support for the Neonatal Research Network's Generic Database Study and Follow-up Study. The authors were supported by grants from the National Institute of Child Health and Human Development and the Department of Health and Human Services (U10 HD21364, U10 HD21373, U10 HD21385, U10 HD21397, U10 HD21415, U10 HD27851, U10 HD27853, U10 HD27856, U10 HD27871, U10 HD27880, U10 HD27881, U10 HD27904, U10 HD34216, U10 HD36790, U10 HD40461, U10 HD40492, U10 HD40498, U10 HD40521, and U10 HD40689) and from the National Institutes of Health (UL1 RR24139, UL1 RR24160, UL1 RR25008, M01 RR30, M01 RR32, M01 RR39, M01 RR44, M01 RR70, M01 RR80, M01 RR125, M01 RR633, M01 RR750, M01 RR997, M01 RR6022, M01 RR7122, M01 RR8084, and M01 RR16587). Funded by the National Institutes of Health (NIH).

Abbreviations

GA Gestational Age

BW	Birth weight
SGA	Small for gestational age
NDI	Neurodevelopmental impairment
AUC	Area under the curve
BSID	Bayley scales of infant development

References

1. Ambalavanan N, Carlo WA, Tyson JE, Langer JC, Walsh MC, Parikh NA, et al. Outcome trajectories in extremely preterm infants. *Pediatrics*. 2012; 130(1):e115–125. [PubMed: 22689874]
2. Ambalavanan N, Carlo WA, Shankaran S, Bann CM, Emrich SL, Higgins RD, et al. Predicting outcomes of neonates diagnosed with hypoxemic-ischemic encephalopathy. *Pediatrics*. 2006; 118(5):2084–2093. [PubMed: 17079582]
3. Medlock S, Ravelli AC, Tamminga P, Mol BW, Abu-Hanna A. Prediction of mortality in very premature infants: a systematic review of prediction models. *PloS one*. 2011; 6(9):e23441. [PubMed: 21931598]
4. Ambalavanan N, Baibergenova A, Carlo WA, Saigal S, Schmidt B, Thorpe KE, et al. Early prediction of poor outcome in extremely low birth weight infants by classification tree analysis. *The Journal of pediatrics*. 2006; 148(4):438–444. [PubMed: 16647401]
5. Ambalavanan N, Carlo WA. Comparison of the prediction of extremely low birth weight neonatal mortality by regression analysis and by neural networks. *Early human development*. 2001; 65(2): 123–137. [PubMed: 11641033]
6. Stoll BJ, Hansen NI, Bell EF, Shankaran S, Laptook AR, Walsh MC, et al. Neonatal outcomes of extremely preterm infants from the NICHD Neonatal Research Network. *Pediatrics*. 2010; 126(3): 443–456. [PubMed: 20732945]
7. Tyson JE, Parikh NA, Langer J, Green C, Higgins RD, National Institute of Child H. et al. Intensive care for extreme prematurity--moving beyond gestational age. *The New England journal of medicine*. 2008; 358(16):1672–1681. [PubMed: 18420500]
8. Leviton A, Blair E, Dammann O, Allred E. The wealth of information conveyed by gestational age. *The Journal of pediatrics*. 2005; 146(1):123–127. [PubMed: 15644836]
9. Lynch CD, Zhang J. The research implications of the selection of a gestational age estimation method. *Paediatric and perinatal epidemiology*. 2007; 21(Suppl 2):86–96. [PubMed: 17803622]
10. Basso O, Wilcox AJ. Might rare factors account for most of the mortality of preterm babies? *Epidemiology*. 2011; 22(3):320–327. [PubMed: 21372718]
11. Cole TJ, Hey E, Richmond S. The PREM score: a graphical tool for predicting survival in very preterm births. *Archives of disease in childhood Fetal and neonatal edition*. 2010; 95(1):F14–19. [PubMed: 19700396]
12. Vohr BR, Wright LL, Poole WK, McDonald SA. Neurodevelopmental outcomes of extremely low birth weight infants <32 weeks' gestation between 1993 and 1998. *Pediatrics*. 2005; 116(3):635–643. [PubMed: 16143580]
13. Vohr BR, Stephens BE, Higgins RD, Bann CM, Hintz SR, Das A, et al. Are Outcomes of Extremely Preterm Infants Improving? Impact of Bayley Assessment on Outcomes. *The Journal of pediatrics*. 2012
14. Fenton TR. A new growth chart for preterm babies: Babson and Benda's chart updated with recent data and a new format. *BMC pediatrics*. 2003; 3:13. [PubMed: 14678563]
15. Hutcheon JA, Zhang X, Platt RW, Cnattingius S, Kramer MS. The case against customised birthweight standards. *Paediatric and perinatal epidemiology*. 2011; 25(1):11–16. [PubMed: 21133965]

16. Vohr BR, Tyson JE, Wright LL, Perritt RL, Li L, Poole WK, et al. Maternal age, multiple birth, and extremely low birth weight infants. *The Journal of pediatrics*. 2009; 154(4):498–503. e492. [PubMed: 19111322]
17. Wadhawan R, Oh W, Perritt RL, McDonald SA, Das A, Poole WK, et al. Twin gestation and neurodevelopmental outcome in extremely low birth weight infants. *Pediatrics*. 2009; 123(2):e220–227. [PubMed: 19139085]
18. Berglund, PA. An introduction to multiple imputation of complex sample data using SAS® v9.2.. SAS Global Forum; 2010.
19. Donovan EF, Tyson JE, Ehrenkranz RA, Verter J, Wright LL, Korones SB, et al. Inaccuracy of Ballard scores before 28 weeks' gestation. National Institute of Child Health and Human Development Neonatal Research Network. *The Journal of pediatrics*. 1999; 135(2 Pt 1):147–152. [PubMed: 10431107]
20. Nguyen TP, Amon E, Al-Hosni M, Gavard JA, Gross G, Myles TD. “Early” versus “late” 23-week infant outcomes. *American journal of obstetrics and gynecology*. 2012; 207(3):226–226. e221–226. [PubMed: 22831809]
21. Anderson PJ, De Luca CR, Hutchinson E, Roberts G, Doyle LW, Victorian Infant Collaborative G. Underestimation of developmental delay by the new Bayley-III Scale. *Archives of pediatrics & adolescent medicine*. 2010; 164(4):352–356. [PubMed: 20368488]
22. Moore T, Johnson S, Haider S, Hennessy E, Marlow N. Relationship between test scores using the second and third editions of the Bayley Scales in extremely preterm children. *The Journal of pediatrics*. 2012; 160(4):553–558. [PubMed: 22048046]
23. Morin I, Morin L, Zhang X, Platt RW, Blondel B, Breart G, et al. Determinants and consequences of discrepancies in menstrual and ultrasonographic gestational age estimates. *BJOG : an international journal of obstetrics and gynaecology*. 2005; 112(2):145–152. [PubMed: 15663577]
24. Wilcox AJ, Weinberg CR, Basso O. On the pitfalls of adjusting for gestational age at birth. *American journal of epidemiology*. 2011; 174(9):1062–1068. [PubMed: 21946386]
25. Ambalavanan N, Carlo WA, Bobashev G, Mathias E, Liu B, Poole K, et al. Prediction of death for extremely low birth weight neonates. *Pediatrics*. 2005; 116(6):1367–1373. [PubMed: 16322160]
26. Wilcox AJ. On the importance--and the unimportance--of birthweight. *International journal of epidemiology*. 2001; 30(6):1233–1241. [PubMed: 11821313]
27. Arnold CC, Kramer MS, Hobbs CA, McLean FH, Usher RH. Very low birth weight: a problematic cohort for epidemiologic studies of very small or immature neonates. *American journal of epidemiology*. 1991; 134:604–13. [PubMed: 1951265]
28. Mongelli M, Wilcox M, Gardosi J. Estimating the date of confinement: ultrasonographic biometry versus certain menstrual dates. *American journal of obstetrics and gynecology*. 1996; 174(1 Pt 1): 278–281. [PubMed: 8572021]
29. Chervenak FA, Skupski DW, Romero R, Myers MK, Smith-Levitin M, Rosenwaks Z, et al. How accurate is fetal biometry in the assessment of fetal age? *American journal of obstetrics and gynecology*. 1998; 178(4):678–687. [PubMed: 9579429]

What is known about this topic: The risk for neurodevelopmental impairment (NDI) and the competing outcome of death in extremely preterm infants is strongly influenced by birth weight (BW) and gestational age (GA). It is unclear which of these two variables is superior for outcome prediction and risk stratification in clinical studies.

What this study adds: Both BW in grams and GA in days are both good predictors of NDI and death in a preterm population selected on the basis of reliable gestational age. Prognostic stratification of baseline risk for NDI or death in extremely preterm infants prior to enrollment into neonatal studies can be accomplished with either BW or GA.

Figure 1. Proportion of infants with NDI or death according to GA and BW

Panel A: Proportion of infants with NDI/death by BW NDI/death rate decreases with increasing BW over all values. **Panel B:** NDI/death rate decreases with increasing GA through 27 weeks (189 days).

Figure 2. Scatter plot of BW by GA to investigate collinearity

Each dot represents 10 infants grouped by similarities in both GA and BW. Values for proportions plotted are raw proportions for each BW and GA interval combination. For combinations with no subjects, occurring mostly at the extremes, the probability plotted is the value of the nearest cell (i.e. BW and GA combination). In most cases, this value is 1.

Table 1

Baseline characteristics

Characteristics		Total
SGA	n	10643
	Yes (%)	762 (7%)
Multiple gestation	n	10652
	Yes (%)	2743 (26%)
Race	n	10569
	Black (%)	4480 (42%)
	White (%)	5637 (53%)
	Other (%)	452 (4%)
Male	n	10652
	Yes (%)	5659 (53%)
Preeclampsia	n	10642
	Yes (%)	1757 (17%)
Any antenatal steroid given prior to delivery to accelerate lung maturity	n	10632
	Yes (%)	7808 (73%)
Gestational Age (day)	n	10652
	Mean (SD)	173 (10)
Birth weight (gram)	N	10650
	Mean (SD)	678 (155)
5-minute Apgar score	N	10506
	Median (Min, Max)	6 (0,10)

Table 2

Comparison of logistic regression models by area under the curve (AUC) at 80% sensitivity for prediction of NDI or death

	Training					Validation	
	AUC (SE)	Specificity	PPV	NPV	Tjur's R^2 ^[1]	AUC (SE)	Tjur's R^2 ^[1]
Unadjusted analysis with BW	0.750 (0.007)	51	78	54	.17	0.750 (0.007)	.17
Unadjusted analysis with GA	0.756 (0.007)	52	79	53	.15	0.755 (0.007)	.15
Unadjusted analysis with BW + GA	0.776 (0.006)	56	80	56	.20	0.775 (0.006)	.19
Unadjusted analysis with BW, GA + Interaction	0.774 (0.006)	55	80	55	.20	0.773 (0.006)	.20
Adjusted analysis without GA or BW	0.745 (0.007)	49	77	53	.17	0.755 (0.007)	.17
Adjusted analysis with BW	0.796 (0.006)	59	81	58	.23	0.802 (0.006)	.24
Adjusted analysis with GA	0.794 (0.006)	57	80	57	.23	0.798 (0.006)	.24
Adjusted analysis with BW + GA	0.802 (0.006)	59	81	58	.24	0.807 (0.006)	.25
Adjusted analysis with BW, GA + Interaction	0.803 (0.006)	61	81	58	.24	0.808 (0.006)	.25

^[1] Tjur's R^2 is an R^2 for logistic models defined as the difference between means of the predicted probabilities of an event for each of the two categories of the dependent variable.

Table 3

Comparison of adjusted logistic regression models by area under the curve (AUC) at 80% sensitivity for prediction of death, NDI, severe IVH (grade 3 or 4), BPD, and PVL.

	AUC (SE)	Specificity	PPV	NPV	p-value
Death through follow-up					
Analysis w/o GA or BW	0.783 (0.006)	55	56	79	0.98
Analysis with BW	0.828 (0.006)	67	64	82	
Analysis with GA	0.828 (0.006)	66	63	82	
Analysis with BW + GA	0.835 (0.005)	68	64	82	
Analysis with BW, GA + Interaction	0.834 (0.005)	67	64	82	
NDI					
Analysis w/o GA or BW	0.627 (0.011)	37	46	73	0.71
Analysis with BW	0.664 (0.011)	41	48	75	
Analysis with GA	0.659 (0.011)	41	48	75	
Analysis with BW + GA	0.668 (0.010)	41	48	76	
Analysis with BW, GA + Interaction	0.668 (0.010)	42	48	76	
BPD or death					
Analysis w/o GA or BW	0.698 (0.007)	41	80	42	0.05
Analysis with BW	0.773 (0.007)	56	84	49	
Analysis with GA	0.755 (0.007)	52	83	47	
Analysis with BW + GA	0.776 (0.006)	57	84	50	
Analysis with BW, GA + Interaction	0.775 (0.006)	56	84	49	
IVH or death					
Analysis w/o GA or BW	0.759 (0.006)	50	61	72	0.45
Analysis with BW	0.798 (0.006)	57	65	74	
Analysis with GA	0.804 (0.006)	61	67	76	
Analysis with BW + GA	0.807 (0.006)	62	67	76	
Analysis with BW, GA + Interaction	0.808 (0.006)	62	67	76	
PVL or death					
Analysis w/o GA or BW	0.774 (0.006)	53	58	77	0.73
Analysis with BW	0.815 (0.006)	63	63	80	
Analysis with GA	0.818 (0.006)	64	64	80	
Analysis with BW + GA	0.823 (0.006)	65	64	80	
Analysis with BW, GA + Interaction	0.823 (0.006)	65	64	80	
BPD					
Analysis w/o GA or BW	0.580 (0.010)	29	61	51	0.02
Analysis with BW	0.662 (0.009)	43	66	61	
Analysis with GA	0.631 (0.009)	35	63	56	
Analysis with BW + GA	0.663 (0.009)	43	66	61	
Analysis with BW, GA + Interaction	0.663 (0.009)	43	66	61	

	AUC (SE)	Specificity	PPV	NPV	p-value
IVH					
Analysis w/o GA or BW	0.637 (0.010)	38	26	88	0.20
Analysis with BW	0.656 (0.010)	39	26	88	
Analysis with GA	0.673 (0.010)	42	27	89	
Analysis with BW + GA	0.673 (0.010)	42	27	89	
Analysis with BW, GA + Interaction	0.673 (0.010)	42	27	89	
PVL					
Analysis w/o GA or BW	0.565 (0.018)	30	6	96	0.88
Analysis with BW	0.566 (0.018)	31	6	96	
Analysis with GA	0.570 (0.018)	31	6	96	
Analysis with BW + GA	0.570 (0.018)	29	6	96	
Analysis with BW, GA + Interaction	0.585 (0.018)	34	6	97	