Neonatal outcomes after maternal biomarker-guided preterm birth intervention: the AVERT PRETERM trial

Matthew K. Hoffman, MD, MPH[1]; Carrie Kitto[1]; Zugui Zhang, PhD[1]; Jing Shi, PhD[2]; Michael G. Walker, PhD[2]; Babak Shahbaba, PhD[3]; Kelly Ruhstaller, MD MSCE[1]

[1]Department of Obstetrics and Gynecology, ChristianaCare, Newark, DE

[2]Walker Bioscience, Carlsbad, CA

[3]Departments of Statistics and Computer Science, University of California Irvine, Irvine, CA

Corresponding Author

Matthew K. Hoffman, MD MPH

Department of Obstetrics & Gynecology

ChristianaCare Health System

4755 Ogletown-Stanton Road

Newark, DE 19718

Telephone: +1 (302) 301-3350

Fax: +1 (302) 733-3340

Email: mhoffman@christianacare.org

Financial Support

Sera Prognostics, Inc.

ClinicalTrials.gov Registration

NCT03151330

Abstract

Objective

To determine whether screening clinically low-risk singleton pregnancies with a validated maternal blood test for preterm birth (PTB) risk, then providing preventive treatments for those at higher risk, will improve neonatal outcomes compared to a large historical population.

Methods

The AVERT PRETERM trial (NCT03151330) took place from June 2018-September 2020 at ChristianaCare Hospital (Newark, DE). Women with singleton non-anomalous pregnancies and no PTB history were prospectively enrolled, screened at $19^{1/7}$- $20^{6/7}$ weeks gestation with a maternal circulating biomarker test for spontaneous PTB (sPTB) risk, and followed through neonatal hospital discharge. Those identified at higher risk (≥16.0%, approximately twice the U.S. population risk) were offered vaginal progesterone (200 mg) and aspirin (81 mg) daily, and care management comprising increased patient outreach and education. Outcomes were compared for prospective arm participants who screened either "not-higher-risk" or "higher-risk accepting treatment" (modified intent-to-treat population) with those in a ChristianaCare historical arm of 10000 pregnancies. Co-primary outcomes were: neonatal morbidity and mortality, as measured by a composite neonatal morbidity index (NMI) score integrating medical complications with neonatal intensive care unit (NICU) length of stay or death; and neonatal hospital length of stay (NNLOS). Cox proportional hazards survival analysis and ordinal logistic regression were used to evaluate outcomes and control for differences between arms.

Results

A COVID-prompted research shutdown occurred when 1460 biomarker-screened individuals had reached 37 weeks' gestation in a racially diverse prospective arm. Of these, 34.7% (507/1460) had biomarker scores indicating higher risk, with 56.4% (286/507) accepting intervention and 43.6% (221/507) declining. The remaining 65.3% (953/1460)

were designated not-higher-risk. Prospective arm neonates had lower NMI scores (odds ratio, 0.81, 95% CI, 0.67-0.98, *P*=0.031), and severe neonatal morbidity (NMI ≥3) was reduced by 18% compared to historical controls. Neonatal hospital stays also were shorter (hazard ratio, 1.35, 95% CI, 1.08-1.70, *P*=0.01), with average NNLOS decreasing by 21% in a quantile of longest stays.

Conclusion

Improved neonatal outcomes after biomarker PTB risk stratification and preventive care management for higher-risk individuals suggests a test-and-treatment strategy that can ameliorate PTB complications and associated medical, societal and economic burdens in a previously unidentifiable patient population: singleton, often nulliparous, pregnancies deemed clinically low risk.

Trial registration

ClinicalTrials.gov #NCT03151330 (https://www.clinicaltrials.gov/study/NCT03151330)

Funding Source

Sera Prognostics, Inc.

Introduction

Preterm birth (PTB) remains the leading cause of perinatal mortality,[1,2] and children born prematurely are at great risk for chronic medical conditions[3,4] and developmental delays. These risks are inversely proportional to the neonate's gestational age at birth (GAB). Survival gains are largely attributable to improved neonatal care[5] and antenatal corticosteroids.[6] Strategies targeting at-risk women, such as vaginal progesterone,[7,8] low-dose aspirin (LDASA),[9] and focused care management[10] comprising increased patient outreach and education may reduce PTB, but their impact has been limited by poor accuracy in identifying at-risk pregnancies. Clinical risk factors include prior PTB[11] and shortened cervical length measured by second-trimester transvaginal ultrasound,[12,13] but their utility is blunted by the fact that most individuals delivering prematurely have not had a prior PTB[14] nor a short cervix at the time of routine sonography (18-22 weeks' gestation).[15,16]

Recent discoveries have identified biomarkers that are differentially expressed in pregnancies delivering prematurely compared to term births.[17] One risk predictor (PreTRM), validated in independent and diverse study populations, measures the ratio of maternal circulating insulin-like growth factor binding protein 4 (IGFBP4) to sex hormone binding globulin (SHBG) in the window of $19^{1/7}$-$20^{6/7}$ weeks' gestation. This predictor stratified risk in U.S. women with an accuracy (area under the receiver operating characteristic curve [AUC]) of 0.80.[18,19] Biological links to spontaneous PTB (sPTB) have been proposed to involve IGFBP4 involvement in sensing fetal nutrient delivery and SHBG involvement in pro-inflammatory signaling within the placenta.[18,20] A risk score threshold corresponding to twice the U.S. population sPTB risk was subsequently validated[21] and shown to significantly stratify higher- and not-higher-risk participants in an extended window of $18^{0/7}$-$20^{6/7}$ weeks' gestation, with a sensitivity and specificity of 88% and 75%, respectively.[19]

The objective of the AVERT PRETERM trial was to test whether targeting treatment improves neonatal outcomes for pregnancies deemed by the biomarker to be at elevated PTB risk but lacking traditional PTB risk factors. Prior studies of the IGFBP4/SHBG predictor have demonstrated that it enriches strongly for severe sPTB and PTB-associated neonatal

health outcomes such as hospital and neonatal intensive care unit (NICU) stays and neonatal morbitidy.[18,20,21] As described previously,[22] to avoid masking the clinical benefit of the test-and-treatment strategy by including the large excess of healthy births not expected to be altered by treatment, primary and secondary hypotheses were tested on quantiles of the population corresponding to the earliest births and the longest hospital and NICU stays.

Methods

Study Design

The AVERT PRETERM trial (NCT03151330) was a prospective cohort study of pregnant individuals compared to historical controls. Co-primary hypotheses posited that PTB risk stratification using the IGFBP4/SHBG test in a clinically low-risk population and focused preventive treatments for those at higher risk will result in: (1) reduced neonatal morbidity or mortality, using a published index;[7] or (2) decreased neonatal hospital lengths of stay (NNLOS) in comparison to historical controls. Co-secondary hypotheses included: (1) reduced NICU length of stay (NICULOS); or (2) prolonged gestational age at birth (GAB) compared to historical controls.

Screening and Recruitment

Prospective and historical control eligibility requirements were singleton pregnancies without evidence of mullerian or fetal anomalies, cervical shortening (<25 mm), genetic anomalies, history of a prior PTB, cervical cerclage, or maternal chronic medical conditions with clear indication for delivery <37 weeks' gestation. Prospective arm participants were identified and screened for eligibility in ambulatory sites and/or at the time of routine ultrasound. Due to requirements of the biomarker test, prospective arm participants were excluded if they had a blood transfusion during the current pregnancy, known hyperbilirubinemia, were taking traditional or low-molecular-weight heparin, or had a known reaction or contraindication to progesterone or aspirin.

Procedures

Following consent, blood was obtained from prospective arm participants within the window of $19^{1/7}$-$20^{6/7}$ weeks' gestation and samples analyzed using the IGFBP4/SHBG test in a Clinical Laboratory Improvement Amendments- and College of American Pathologists-certified laboratory (Sera Prognostics, Inc., Salt Lake City, UT), as detailed previously.[18,23]

Test results were shared with the participant and their care provider. Those with a risk score ≥16.0% (approximately twice the U.S. population PTB risk) were designated as higher-risk then offered and consented again to receive care management, consisting of twice-weekly nursing contacts to monitor medication compliance and symptom development, and daily progesterone (200 mg intravaginally) and aspirin (81 mg). The remainder of care was determined by the treating clinician. Individuals with a risk score <16.0% were designated as not-higher-risk and received usual obstetric care.

Outcomes for both arms were obtained through a validated obstetrical registry[24] or, for prospective arm participants who delivered elsewhere, through medical record review. External data review was performed for the prospective arm and 10% of historical controls to ensure that eligibility requirements were met, with an eligibility error rate of <3% deemed acceptable in the historical arm. All PTB cases were reviewed further for accurate assessment of primary outcomes by a single investigator (MKH). As secular changes in care over the study period could affect prospective arm outcomes in a non-random fashion, major changes in guidance or management protocols were documented on a quarterly basis.

Trial Outcomes

Two co-primary outcomes were selected: (1) NMI, measured using a composite ordinal index score, "NMI with neonatal intensive care unit (NICU) length of stay (NICULOS)",[7] which also factored in neonatal death. (2) NNLOS from time of birth to discharge. Medical complications included in the NMI and scoring on a severity scale of 0 to 4, with scores ≥3 indicating severe morbidity and 4 indicating neonatal death, are detailed in **Appendix 1**. Co-secondary outcomes were GAB and NICULOS. These endpoints are better-powered

continuous health-related outcomes than typical dichotomous surrogate endpoints (e.g., sPTB rate)[22,25] and have been reported to be enriched by the biomarker test.[20,21] Since the competing risk of death results in shorter stays, the protocol prespecified that NNLOS and NICULOS for fetuses or neonates who expired would be adjusted to be one day longer than the longest stay observed among all neonates. To avoid bias, calculations of mean days saved in the NICU between arms used the actual NICULOS for any neonates that expired in the NICU.

Power and Sample Size Estimation

Based on historical pregnancy data from ChristianaCare, it was estimated that 10000 consecutive historical controls from an approximately two-year period immediately prior to study initiation would be available. The required sample size for the prospective arm was then determined assuming 10000 control subjects. From historical pregnancy data at ChristianaCare, a PTB rate of 9.1% was estimated. Sample size estimation was built from the co-primary outcomes using a simulated GAB distribution with a singleton PTB rate of 9.1% and an effect of interventions described elsewhere.[26,27] α-level spending of 0.05 was shared between two outcomes using Holm's method.[28]

For the NMI co-primary outcome, power was estimated conservatively by a binary comparison of the proportion of subjects with NMI scores ≥3, assumed to be 2.0-2.3% in the prospective arm and near 3.6% in the historical arm, based on a previous clinical utility study.[25] Assuming these proportions, with 55% compliance among higher-risk subjects, and approximately 10000 historical controls, a Fisher Exact test with a sample size of approximately 1453 subjects with outcomes in the prospective arm would provide power of 0.7-0.9.[29]

For the NNLOS co-primary outcome, the hazard ratio (HR) was expected to be 1.32-1.46, based on simulations using data from a previous clinical utility study.[25] Assuming these HRs, with 55% compliance among higher-risk subjects and approximately 10000 historical controls, a Cox proportional hazards (PH) analysis with a sample size of

approximately 1453 subjects with outcomes in the prospective arm would provide power of at least 0.8.[29]

Statistical Analysis

Variables were summarized using means and standard deviations for continuous data, and percentages and frequencies for categorical data. Comparisons for baseline characteristics were performed; the Wilcoxon rank sum test was used to compare continuous variables between two groups, and contingency table analysis (chi-square) was used to compare categorical variables, with significance set to $P<0.05$. During study design, it was recognized that care for newborns with neonatal opioid withdrawal syndrome (NOWS) would differ between the prospective and the historical arms due to implementation of the Eat, Sleep, Console approach.[30] Prespecified covariates included in the primary and secondary analysis models included maternal age, parity and maternal opioid use, assessed as NOWS. Additional covariates were examined using sensitivity analyses.

The NNLOS co-primary hypothesis and the NICULOS and GAB co-secondary hypotheses were tested using Cox PH regression, adjusted for covariates. The NMI co-primary hypothesis was tested using ordinal logistic regression, adjusted for covariates.

Co-primary and co-secondary analyses used a modified intent-to-treat (mITT) population, defined as all subjects for whom both co-primary outcomes were known; and who were either selected for the historical arm, received a not-higher-risk biomarker test result, or a higher-risk biomarker test result with consent and treatment initiation with vaginal progesterone and LDASA before 24 weeks' gestation.

As prespecified by a defined selection procedure, subjects in the 8.5% quantile of longest stays were included in the NICULOS and NNLOS analyses, while subjects in the earliest 8.5% quantile were included in the GAB analysis. Specifically, the prespecified quantile was 1.2 times the observed PTB rate in the historical arm (7.1%). This approach captures the impact of interventions on severe outcomes, since interventions were not expected to affect the vast excess of healthy babies delivered full term. Using a quantile

slightly larger than the PTB rate can capture the full benefit incurred by PTB prevention.[22] Prespecified exploratory analyses included evaluating outcomes on the full mITT and intent-to-treat (ITT) populations.

Prespecified exploratory analyses included testing outcomes in the full mITT and ITT populations. To examine the effect size at binary NMI cutoffs, differences between arms at each NMI level were calculated using a logistic regression model with a binary response variable and the covariates maternal age, parity and maternal opioid use. Percent differences between arms for time-to-event analyses were calculated using the hazard rate for each group as estimated by the fitted Cox PH model with the same covariates.

Analyses were performed using R software version 4.2.2.[31] The brant function from the brant package was used to test the proportional odds assumption, the cox.zph function from the survival package was used to test the PH assumption, and the forestplot package was used to generate forest plots. *P*-values <0.05 were considered significant. The co-primary outcome analyses used Holm's multiple comparisons correction.[28]

Trial Oversight

The study protocol was approved by the ChristianaCare institutional review board prior to participant enrollment. An independent data and safety monitoring board convened prior to study initiation, approved the protocol, and provided oversight of adverse events. All participants provided written informed consent, and all authors accept responsibility for the accuracy and completeness of the data and for fidelity in the conduct of the trial.

Results

The AVERT PRETERM prospective arm was enrolled and followed from June 2018-September 2020 at ChristianaCare Hospital (Newark, DE), a regional health care system that serves a mixed urban and rural population across Delaware and Maryland. The historical control arm delivered at ChristianaCare from August 2016-July 2018. In April 2020, ChristianaCare Health System halted all non-COVID research. The study was terminated and the statistical analysis plan reassessed in a blinded manner. A 2020 study[32]

showed that COVID-19 infection led to an increase in stillbirths. To avoid bias in comparing pre-pandemic controls to prospective subjects who reached term during the pandemic period, the plan was modified to limit the primary analysis to subjects who had reached 37 weeks' gestation before the local spread of SARS-CoV-2. Over the prospective and historical periods of the study, no major changes were made to study or neonate care protocols, other than as noted for NOWS.

At study termination, 1873 eligible subjects had been enrolled in the prospective arm, 1460 of whom were eligible and aligned with pre-COVID-19 patient care conditions and had been screened with the biomarker test (**Figure 1**). Of these, 34.7% (507/1460) were deemed higher-risk by the test. Among screened individuals, 83.4% (1218/1460) had clinical outcomes and qualified either as not-higher-risk (77.1%, 939/1218) or as higher-risk accepting treatment (22.9%, 279/1218). Prospective arm outcomes were compared to those for 10000 consecutive historical controls selected from an approximately two-year period immediately preceding study initiation.

Baseline participant characteristics and delivery data are shown in **Table 1**. The prospective arm was significantly older, more obese, and more likely to be multiparous, have hypertension and smoke than historical controls. Similarly, body mass index was higher in the prospective arm compared to historical controls – mostly due to higher weight, though the prospective arm was nominally taller than historical controls. The proportion of Black participants in both arms was 26.5%.

Results of hypothesis tests for co-primary and co-secondary outcomes are presented in **Table 2**. NMI scores were significantly reduced in the prospective arm versus the historical arm (OR 0.81, 95% CI 0.67-0.98, P=0.03) (**Table 2**). Specifically, the probability of NMI ≥1 (any impairment) or NMI ≥3 (severe NMI) was reduced on average by 13%-18%, respectively, across a range of covariate values (**Appendix 2**). NNLOS was significantly reduced in the prospective arm versus the historical arm (HR 1.35, 95% CI, 1.08-1.7, P=0.01). The Kaplan-Meier plot for NNLOS is shown in **Figure 2**, reflecting a 21% reduction in mean NNLOS (Cox PH P=0.01) in a quantile of longest stays. Neither the

proportional odds assumption for the intervention effects in the ordinal logistic regression nor the PH assumption for the prospective arm in the Cox regression were violated.

After statistical significance was achieved for both co-primary outcomes, the co-secondary hypotheses were tested. In Cox regression analysis, neonates tended to leave the NICU earlier (HR 1.2, 95% CI, 0.96-1.51), but this difference was not significant (*P*=0.12). No difference in GAB was observed (HR 0.962, 95% CI, 0.84-1.10, *P*=0.584).

In a non-prespecified analysis, although no GAB difference was seen between arms as tested in the secondary analyses, the number of individuals delivering <32 weeks' gestation decreased (HR, 0.52, 95%CI, 0.28-0.94, Cox PH *P*=0.009) (**Figure 3a**), corresponding to a 2.5-week extension of mean gestation (29.93 and 27.46 weeks for prospective and historical arms, respectively). Prospective arm neonates born <32 weeks' gestation left the hospital earlier than did those in the historical arm (HR, 1.84, 95% CI 1.01-3.36, Cox PH *P*=0.046), with mean NNLOS differences of approximately 30% (mean of 68.47 and 97.23 days for prospective and historical arms, respectively) (**Figure 3b**).

For prespecified exploratory analyses, outcomes are reported for all subjects in the mITT and ITT populations. Neonates were discharged from the hospital (NNLOS) and the NICU (NICULOS) earlier in the prospective arm than in historical controls (**Figure 4a**). Mean observed NICULOS savings across all pregnancies tested corresponded to 0.55 days (95%CI, 0.018-1.078, *P*=0.043) in the mITT population and 0.60 days (95%CI, 0.107-1.087, *P*=0.017) in the ITT population. GAB <32 weeks' gestation was significantly prolonged in the ITT population (*P*=0.045) (**Figure 4a**). Mean GAB across all mITT participants was 38.6 and 38.5 weeks for the historical and prospective arms, respectively, and 38.6 and 38.4 weeks in ITT. Odds ratios between arms for PTB and sPTB at selected GAB cutoffs are shown in **Figure 4b**. While PTB and sPTB <37 weeks' gestation do not differ significantly between arms, those for <35, <32, and <28 week cutoffs trend lower in both the mITT and ITT populations. NMI, significant in the mITT primary analysis, remained favorably reduced in the ITT population.

For those identified as higher-risk, biomarker test scores were significantly higher in the group that accepted intervention than in the group that did not. This difference made

comparison of outcomes for pregnancies deemed higher-risk accepting treatment and not accepting treatment challenging to interpret.

Discussion

Evidence exists for intervention benefit among individuals stratified at higher PTB risk using clinical factors.[7-10] By extension, it is reasonable to postulate that biomarker-based PTB risk stratification of otherwise low-risk pregnancies can identify individuals who might benefit from bundled interventions. AVERT PRETERM trial results demonstrate that biomarker-based PTB risk screening and treatment with care management, LDASA, and vaginal progesterone, resulted in deliveries with shorter neonatal hospital stays and less severe neonatal morbidities compared to a large historical arm, after controlling for population differences. Clinically, the impact appears greatest amongst pregnancies delivering <32 weeks' gestation, which remains the primary driver of newborn and child morbidity and mortality.

These findings suggest a strategy for universal screening and treatment to ameliorate PTB complications amongst women lacking traditional clinical risk factors. The results resonate with those from a randomized investigation[25] of 1191 women who either received biomarker test results and treatment or did not receive results. In that study, the NICULOS due to sPTB was significantly shorter amongst those screened and treated versus those not receiving results (median 6.8 days versus 45.5 days; P=0.005). Evidence for a GAB shift has been reported for that study,[22] though it was limited by low sample numbers. Moreover, a randomized controlled trial evaluating the clinical utility of the biomarker test (ClinicalTrials.gov #NCT04301518) was recently stopped at interim analysis due to efficacy, and reporting of results is pending.

Given the substantially higher PTB rate among U.S. Black women (14.4%) as compared to that across the entire population (10.2%),[2] it is important to note that Black participants were represented in both AVERT PRETERM study arms with a proportion (26.5%) nearly double recent population estimates (13.6%).[2] Results in this study population, along with those in two large and similarly diverse studies,[18,20] indicate that the

biomarker test will be applicable across the diverse U.S. population. Establishing validated PTB risk with a validated biomarker test enables implementation of effective PTB prevention strategies among clinically low-risk individuals and, potentially, improvement in patient adherence to usual prenatal care.

Strengths and Limitations

Study strengths include a multimodal intervention strategy to mitigate complications of premature delivery, as well as a biomarker test that has been validated in multiple cohorts. The data registry[24] used to obtain historical control data from the EMR has been well validated, medical record review of subjects delivering at other institutions was conducted, and all PTB cases were further reviewed by a single investigator.

Study limitations include imbalances inherent in comparison of a prospective arm with a historical arm that differed in several maternal demographic and medical conditions. These differences were addressed through multivariable modeling but remain a potential source of bias. Even so, significant demographic differences in the prospective versus historical arm – older age, more hypertension and more smoking – likely biases the prospective arm toward increased PTB incidence, further underscoring the importance of these findings. Additionally, for LDASA, guideline changes have expanded the number of women eligible for treatment. A recent estimate suggests that most pregnant people should be counseled about LDASA.[33] One may argue that both progesterone and care management have limited evidence of efficacy in the situations studied here. However, the multifactorial etiologies of PTB make proving that either or both are effective in a broad range of individuals difficult, as an intervention may work on some etiologies but not others. The value of using this suite of interventions with clinically low-risk pregnancies, at least in theory and perhaps supported by results herein, is the potential for broader coverage of etiologies and better hope of success through multiple interventions. Finally, there is overlap among trial outcomes, as NICULOS was included in the NMI index, and no multiple comparison adjustment of *P* values was made for co-secondary and exploratory outcomes.

Conclusions

Screening singleton, non-anomalous pregnancies lacking traditional clinical risk factors with a validated biomarker blood test for PTB risk prediction, then targeting preventive treatments for those with higher risk, shortened neonatal hospital stays and reduced neonatal morbidity. This test-and-treatment strategy can ameliorate PTB complications and associated medical, societal, and economic burdens in a large yet previously unidentifiable population: singleton, often nulliparous, pregnancies deemed clinically low risk.

Acknowledgments

Sera Prognostics, Inc. provided funding to conduct this independent investigation. Jennifer Logan, PhD, an employee of Sera Prognostics, Inc., supported manuscript preparation. A preprint of this manuscript is posted at https://doi.org/10.1101/2023.09.13.23295503

Conflict of Interest Disclosure

BS, JS, and MGW are paid consultants of Sera Prognostics, Inc. For the purposes of this study, JS and MGW reported to MKH and were paid consultants of ChristianaCare, and BS was supported by Sera Prognostics, Inc. MKH received an investigator-initiated grant from Sera Prognostics, Inc., to conduct this study.

References

1. Liu L, Oza S, Hogan D, Chu Y, Perin J, Zhu J, et al. Global, regional, and national causes of under-5 mortality in 2000-15: an updated systematic analysis with implications for the Sustainable Development Goals. *Lancet*. Dec 17 2016;388(10063):3027-3035. doi:10.1016/S0140-6736(16)31593-8

2. Martin JA, Hamilton BE, Osterman MJK. Births in the United States, 2022. *NCHS Data Brief*. Aug 2023;(477):1-8.

3. Moster D, Lie RT, Markestad T. Long-term medical and social consequences of preterm birth. *N Engl J Med*. Jul 17 2008;359(3):262-73. doi:10.1056/NEJMoa0706475

4. Crump C, Sundquist K, Sundquist J, Winkleby MA. Gestational age at birth and mortality in young adulthood. *JAMA*. Sep 21 2011;306(11):1233-40. doi:10.1001/jama.2011.1331

5. Richardson DK, Gray JE, Gortmaker SL, Goldmann DA, Pursley DM, McCormick MC. Declining severity adjusted mortality: evidence of improving neonatal intensive care. *Pediatrics*. Oct 1998;102(4 Pt 1):893-9. doi:10.1542/peds.102.4.893

6. Crowley PA. Antenatal corticosteroid therapy: a meta-analysis of the randomized trials, 1972 to 1994. *Am J Obstet Gynecol*. Jul 1995;173(1):322-35. doi:10.1016/0002-9378(95)90222-8

7. Hassan SS, Romero R, Vidyadhari D, Fusey S, Baxter JK, Khandelwal M, et al. Vaginal progesterone reduces the rate of preterm birth in women with a sonographic short cervix: a multicenter, randomized, double-blind, placebo-controlled trial. *Ultrasound Obstet Gynecol*. Jul 2011;38(1):18-31. doi:10.1002/uog.9017

8. Fonseca EB, Celik E, Parra M, Singh M, Nicolaides KH, Fetal Medicine Foundation Second Trimester Screening G. Progesterone and the risk of preterm birth among women with a short cervix. *N Engl J Med*. Aug 2 2007;357(5):462-9. doi:10.1056/NEJMoa067815

9. Hoffman MK, Goudar SS, Kodkany BS, Metgud M, Somannavar M, Okitawutshu J, et al. Low-dose aspirin for the prevention of preterm delivery in nulliparous women with a singleton pregnancy (ASPIRIN): a randomised, double-blind, placebo-controlled trial. *Lancet*. Jan 25 2020;395(10220):285-293. doi:10.1016/S0140-6736(19)32973-3

10. Garite TJ, Manuck TA. Should case management be considered a component of obstetrical interventions for pregnancies at risk of preterm birth? *Am J Obstet Gynecol*. Apr 2023;228(4):430-437. doi:10.1016/j.ajog.2022.09.022

11. Mercer BM, Goldenberg RL, Moawad AH, Meis PJ, Iams JD, Das AF, et al. The preterm prediction study: effect of gestational age and cause of preterm birth on subsequent obstetric outcome. National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network. *Am J Obstet Gynecol*. Nov 1999;181(5 Pt 1):1216-21. doi:10.1016/s0002-9378(99)70111-0

12. Iams JD, Goldenberg RL, Meis PJ, Mercer BM, Moawad A, Das A, et al. The length of the cervix and the risk of spontaneous premature delivery. National Institute of Child Health and Human Development Maternal Fetal Medicine Unit Network. *N Engl J Med*. Feb 29 1996;334(9):567-72. doi:10.1056/NEJM199602293340904

13. Cahill AG, Odibo AO, Caughey AB, Stamilio DM, Hassan SS, Macones GA, Romero R. Universal cervical length screening and treatment with vaginal progesterone to prevent preterm birth: a decision and economic analysis. *Am J Obstet Gynecol*. Jun 2010;202(6):548 e1-8. doi:10.1016/j.ajog.2009.12.005

14. Petrini JR, Callaghan WM, Klebanoff M, Green NS, Lackritz EM, Howse JL, et al. Estimated effect of 17 alpha-hydroxyprogesterone caproate on preterm birth in the United States. *Obstet Gynecol*. Feb 2005;105(2):267-72. doi:10.1097/01.AOG.0000150560.24297.4f

15. Esplin MS, Elovitz MA, Iams JD, Parker CB, Wapner RJ, Grobman WA, et al. Predictive Accuracy of Serial Transvaginal Cervical Lengths and Quantitative Vaginal Fetal Fibronectin Levels for Spontaneous Preterm Birth Among Nulliparous Women. *JAMA*. Mar 14 2017;317(10):1047-1056. doi:10.1001/jama.2017.1373

16. Berghella V, Saccone G. Cervical assessment by ultrasound for preventing preterm delivery. *Cochrane Database Syst Rev*. Sep 25 2019;9(9):CD007235. doi:10.1002/14651858.CD007235.pub4

17. Lamont RF, Richardson LS, Boniface JJ, Cobo T, Exner MM, Christensen IB, et al. Commentary on a combined approach to the problem of developing biomarkers for the

prediction of spontaneous preterm labor that leads to preterm birth. *Placenta*. Sep 1 2020;98:13-23. doi:10.1016/j.placenta.2020.05.007

18. Saade GR, Boggess KA, Sullivan SA, Markenson GR, Iams JD, Coonrod DV, et al. Development and validation of a spontaneous preterm delivery predictor in asymptomatic women. *Am J Obstet Gynecol*. May 2016;214(5):633 e1-633 e24. doi:10.1016/j.ajog.2016.02.001

19. Burchard J, Saade GR, Boggess KA, Markenson GR, Iams JD, Coonrod DV, et al. Better Estimation of Spontaneous Preterm Birth Prediction Performance through Improved Gestational Age Dating. *J Clin Med*. May 19 2022;11(10)doi:10.3390/jcm11102885

20. Markenson GR, Saade GR, Laurent LC, Heyborne KD, Coonrod DV, Schoen CN, et al. Performance of a proteomic preterm delivery predictor in a large independent prospective cohort. *Am J Obstet Gynecol MFM*. Aug 2020;2(3):100140. doi:10.1016/j.ajogmf.2020.100140

21. Burchard J, Polpitiya AD, Fox AC, Randolph TL, Fleischer TC, Dufford MT, et al. Clinical Validation of a Proteomic Biomarker Threshold for Increased Risk of Spontaneous Preterm Birth and Associated Clinical Outcomes: A Replication Study. *J Clin Med*. Oct 29 2021;10(21):5088. doi:10.3390/jcm10215088

22. Combs CA, Zupancic JAF, Walker M, Shi J. Prediction and Prevention of Preterm Birth: Secondary Analysis of a Randomized Intervention Trial. *J Clin Med*. Aug 23 2023;12(17)doi:10.3390/jcm12175459

23. Bradford C, Severinsen R, Pugmire T, Rasmussen M, Stoddard K, Uemura Y, et al. Analytical validation of protein biomarkers for risk of spontaneous preterm birth. *Clinical Mass Spectrometry*. 2017/01/01/ 2017;3:25-38. doi:10.1016/j.clinms.2017.06.002

24. Zhang J, Troendle J, Reddy UM, Laughon SK, Branch DW, Burkman R, et al. Contemporary cesarean delivery practice in the United States. *Am J Obstet Gynecol*. Oct 2010;203(4):326 e1-326 e10. doi:10.1016/j.ajog.2010.06.058

25. Branch DW, VanBuren JM, Porter TF, Holmgren C, Holubkov R, Page K, et al. Prediction and Prevention of Preterm Birth: A Prospective, Randomized Intervention Trial. *Am J Perinatol*. Jul 2023;40(10):1071-1080. doi:10.1055/s-0041-1732339

26. Grabner M, Burchard J, Nguyen C, Chung H, Gangan N, Boniface JJ, et al. Cost-Effectiveness of a Proteomic Test for Preterm Birth Prediction. *Clinicoecon Outcomes Res*. 2021;13:809-820. doi:10.2147/CEOR.S325094

27. Burchard J, Markenson GR, Saade GR, Laurent LC, Heyborne KD, Coonrod DV, et al. Clinical and economic evaluation of a proteomic biomarker preterm birth risk predictor: cost-effectiveness modeling of prenatal interventions applied to predicted higher-risk pregnancies within a large and diverse cohort. *J Med Econ*. Jan-Dec 2022;25(1):1255-1266. doi:10.1080/13696998.2022.2147771

28. Holm S. A Simple Sequentially Rejective Multiple Test Procedure. *Scand J Stat*. 1979;6(2):65-70.

29. NCSS Statistical Software. Group-Sequential Analysis for Two Proportions. Accessed May 1, 2023, https://www.ncss.com/wp-content/themes/ncss/pdf/Procedures/NCSS/Group-Sequential_Analysis_for_Two_Proportions.pdf

30. Young LW, Ounpraseuth ST, Merhar SL, Hu Z, Simon AE, Bremer AA, et al. Eat, Sleep, Console Approach or Usual Care for Neonatal Opioid Withdrawal. *N Engl J Med*. Jun 22 2023;388(25):2326-2337. doi:10.1056/NEJMoa2214470

31. R Foundation. The R Project for Statistical Computing. Accessed August 1, 2022, https://www.r-project.org/

32. Khalil A, von Dadelszen P, Draycott T, Ugwumadu A, O'Brien P, Magee L. Change in the Incidence of Stillbirth and Preterm Delivery During the COVID-19 Pandemic. *JAMA*. Jul 10 2020;324(7):705-6. doi:10.1001/jama.2020.12746

33. Wheeler SM, Myers SO, Swamy GK, Myers ER. Estimated Prevalence of Risk Factors for Preeclampsia Among Individuals Giving Birth in the US in 2019. *JAMA Netw Open*. Jan 4 2022;5(1):e2142343. doi:10.1001/jamanetworkopen.2021.42343

Table 1. Baseline characteristics and delivery data in the modified intent-to-treat population.

	Historical arm (N = 10000)	Prospective arm (N = 1218)	*P*
Maternal Age			<0.001
N	10000	1218	
Mean (SD)	29.6 (5.4)	30.5 (5.7)	
Gravida			<0.001
N	9954	1135	
Mean (SD)	2.7 (1.7)	2.41 (1.5)	
Parity			<0.001
N	9953	1159	
Mean (SD)	1.1 (1.2)	0.9 (1.1)	
Percent nulliparous [N, (%)]	6544 (65.7)	630 (54.4)	<0.001
Number of miscarriages			0.25
N	9953	1130	
Mean (SD)	1.6 (1.1)	1.6 (1.0)	
Race* [N, (%)]			<0.001
American Indian	21 (0.2)	1 (0.1)	
Asian	783 (7.8)	76 (6.3)	
Black	2653 (26.5)	322 (26.5)	
White	5634 (56.3)	740 (61.0)	
Other	909 (9.1)	74 (6.1)	
Prepregnancy BMI			0.04
N	9476	728	
Mean (SD)	27.5 (8.5)	28.2 (7.6)	
BMI <19 kg/M^2 [N, (%)]	403 (4.3)	32 (4.4)	0.85
Height (inches)			<0.001
N	9838	1033	
Mean (SD)	64.1 (2.7)	64.48 (2.69)	
Diabetes [N, (%)]	127 (1.3)	19 (1.6)	0.42
Opioid Use [N, (%)]	242 (2.4)	13 (1.1)	0
Hypertension [N, (%)]	606 (6.1)	105 (8.6)	<0.001
Smoking [N, (%)]	709 (7.8)	100 (9.5)	0.06
Insurance type [N, (%)]			0.67
Government	2969 (29.7)	315 (28.6)	
Other	19 (0.2)	1 (0.1)	
Private	7012 (70.1)	787 (71.4)	
Delivery type [N, (%)]			<0.001
Dilation & evacuation	0 (0.0)	1 (0.1)	
Primary cesarean delivery	1577 (15.8)	283 (20.2)	
Repeat cesarean delivery	1541 (15.4)	177 (12.6)	
Vaginal delivery	6630 (66.3)	923 (65.8)	
Vaginal delivery after cesarean	252 (2.5)	18 (1.3)	

BMI, body mass index.
*Race is based on self-report.

Table 2: Results of hypothesis tests for the co-primary and co-secondary outcomes using prespecified covariates in the upper 8.5% quantile of the modified intent-to-treat population.

Co-primary endpoints				
Neonatal length of hospital stay (days) [§]	**Hazard ratio**	**Lower confidence limit**	**Upper confidence limit**	***P*** [‡]
Prospective vs historical arm	1.35	1.08	1.70	0.01
Maternal age	1.00	0.99	1.01	0.75
Parous or nulliparous	0.94	0.82	1.08	0.39
Opioid use	1.53	1.30	1.79	<0.001
Neonatal morbidity index score [¶]	**Odds ratio**	**Lower confidence limit**	**Upper confidence limit**	***P*** [‡]
Prospective vs historical arm	0.81	0.67	0.98	0.03
Maternal age	1.01	1.00	1.02	0.13
Parous or nulliparous	0.58	0.52	0.65	<0.001
Opioid use	2.62	1.97	3.45	<0.001
Co-secondary endpoints				
NICU length of stay (days)* [§]	**Hazard ratio**	**Lower confidence limit**	**Upper confidence limit**	***P*** [‡]
Prospective vs historical arm	1.20	0.96	1.51	0.12
Maternal age	1.00	0.99	1.01	0.78
Parous or nulliparous	0.96	0.84	1.10	0.57
Opioid use	2.36	1.71	3.27	<0.001
Gestational age at birth (weeks) [†§]	**Hazard ratio**	**Lower confidence limit**	**Upper confidence limit**	***P*** [‡]
Prospective vs historical arm	0.96	0.84	1.10	0.58
Maternal age	1.00	0.99	1.01	0.93
Parous or nulliparous	0.85	0.77	0.93	<0.001
Opioid use	0.91	0.73	1.14	0.41

GAB, gestational age at birth; NICU, neonatal intensive care unit; NMI, neonatal morbidity and mortality index score; NICULOS, NICU length of stay; NNLOS, neonatal hospital length of stay.

*NNLOS and NICULOS are reported for subjects in the 8.5% quantile of longest stays in each arm.

[†]GAB is reported for the earliest 8.5% quantile of births in each arm.

[‡]P- values report the significance of the individual covariates listed, and outcome comparisons between the prospective and historical arms are adjusted for these covariates.

[§]Cox proportional hazards regression analysis.

[¶]Ordinal logistic regression analysis.

Figure Legends

Figure 1. Consort diagram for study inclusion.

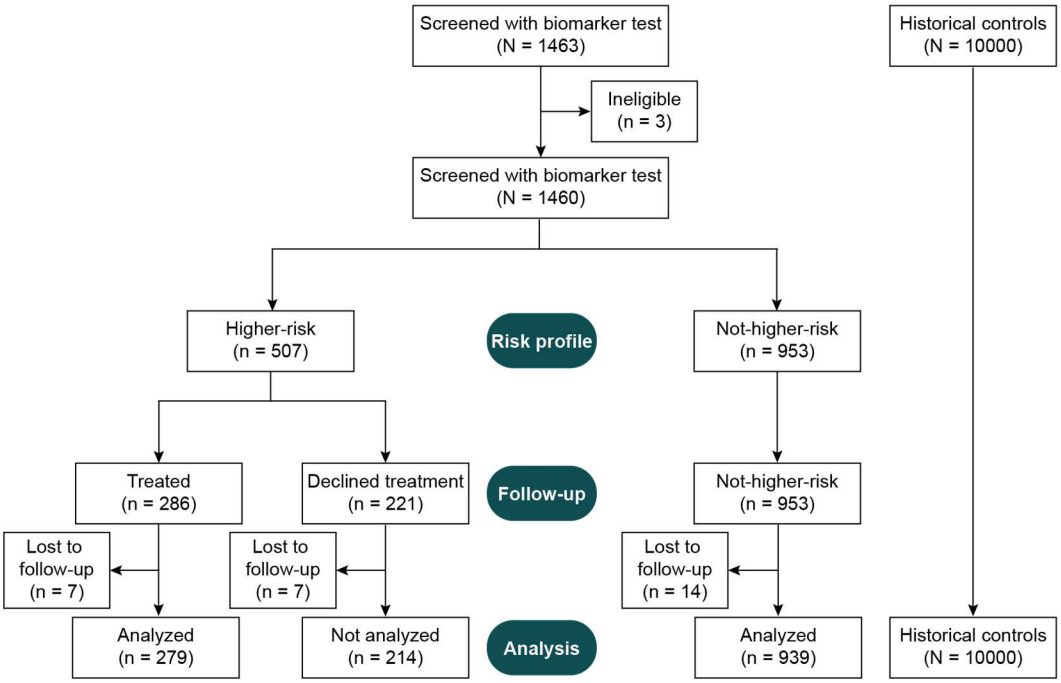


Figure 2: Kaplan-Meier analysis of neonatal hospital length of stay (NNLOS). *P* value was calculated using Cox proportional hazards (PH) regression analysis with covariates.

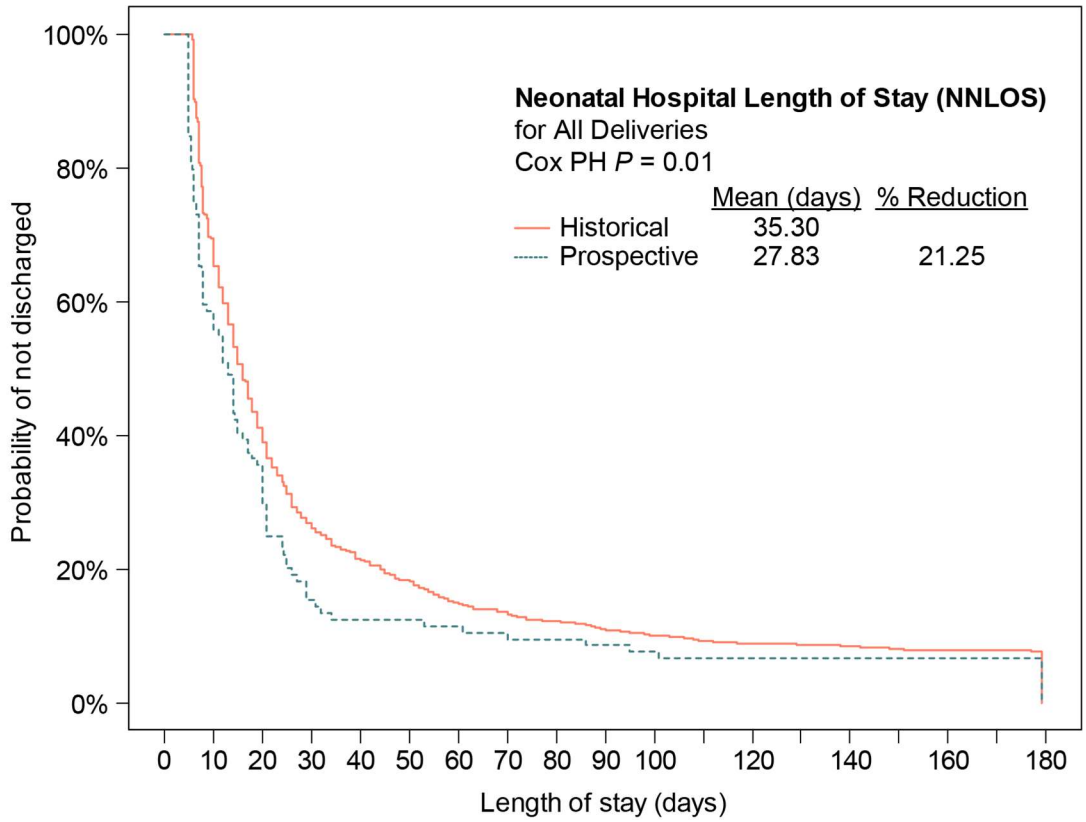


Figure 3: Kaplan-Meier analyses for (A) gestational age at birth (GAB) and (B) neonatal hospital length of stay (NNLOS) for neonates delivered <32 weeks' gestation. *P* values were calculated using Cox proportional hazards (PH) regression analysis with covariates.

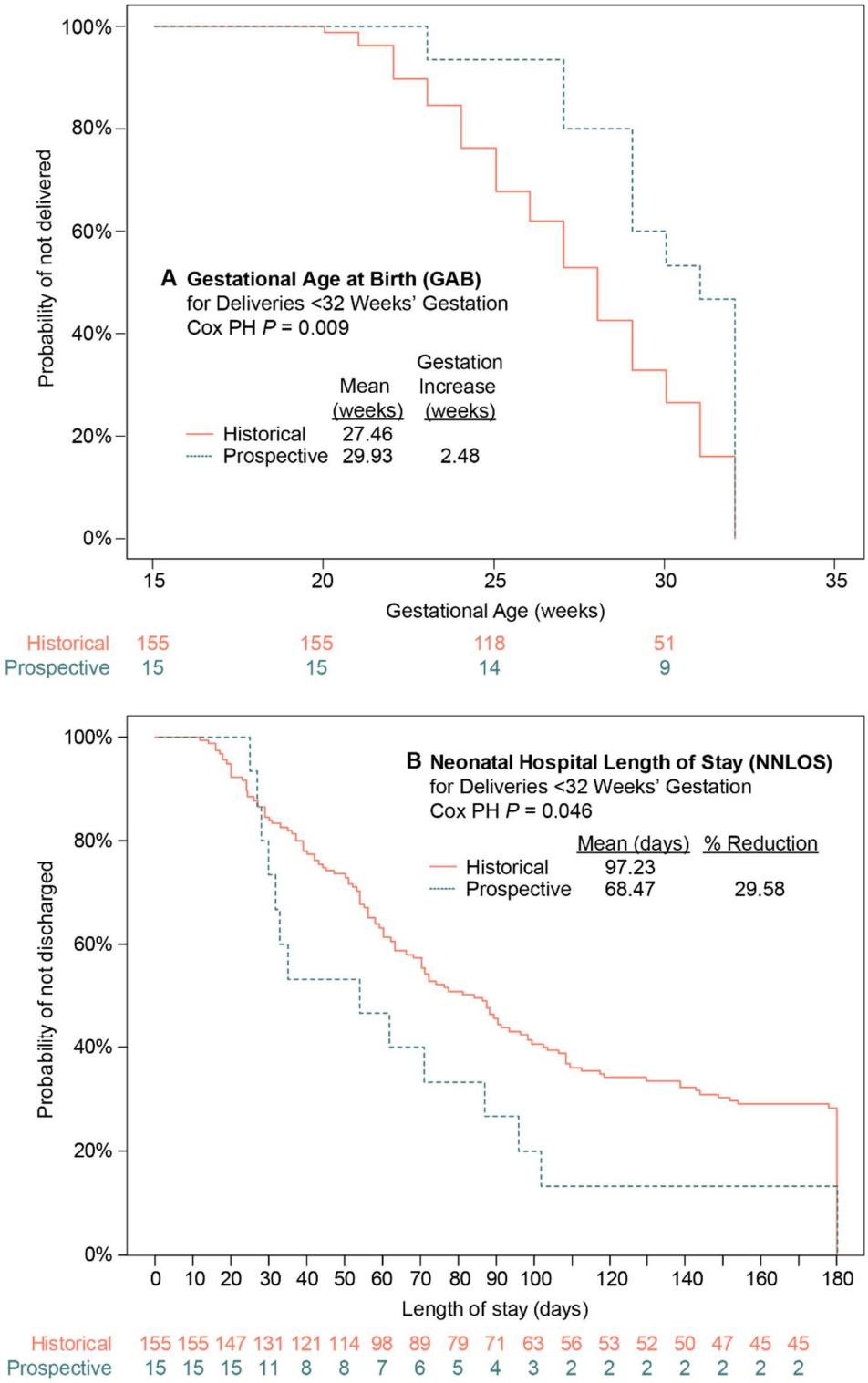
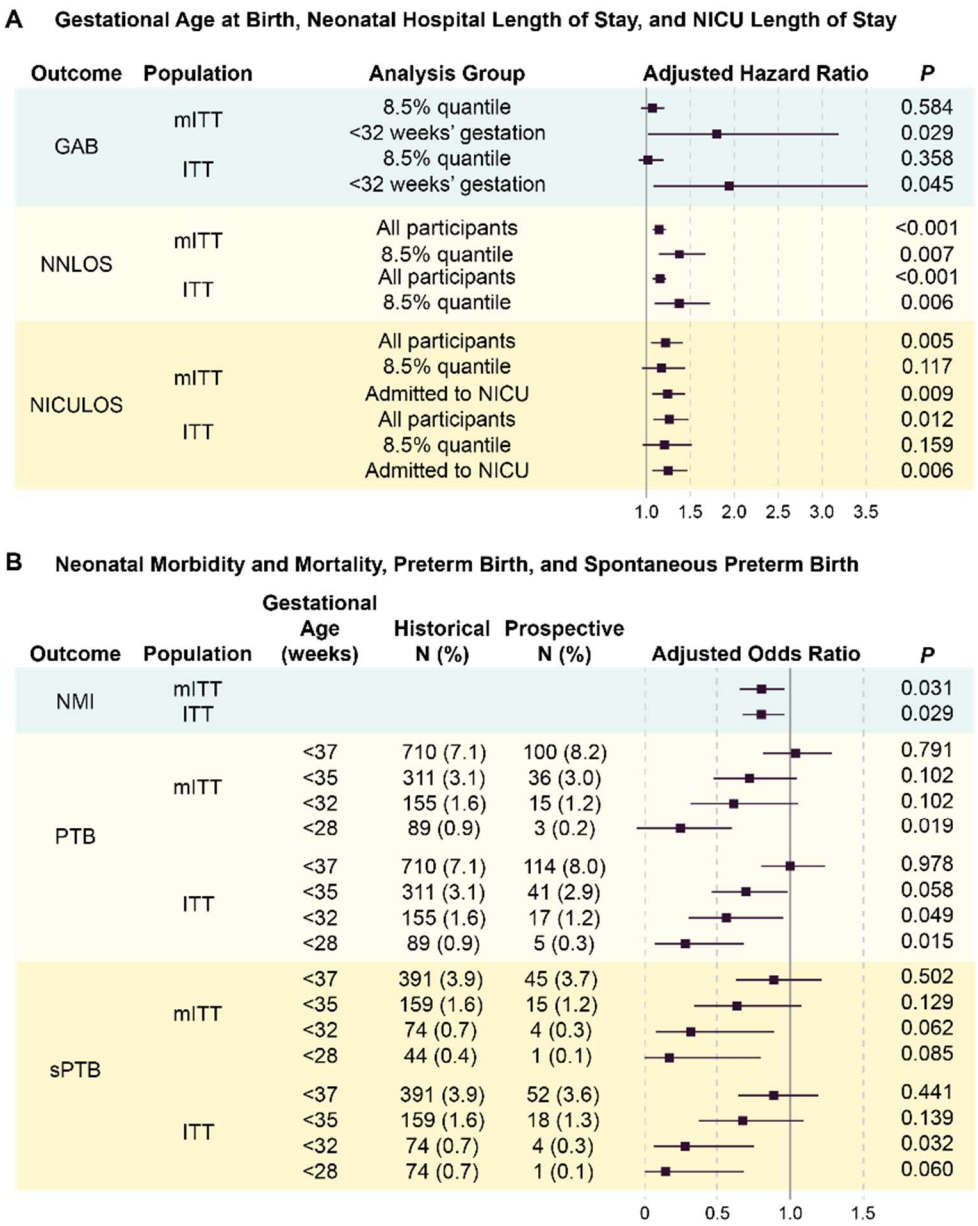


Figure 4. Prespecified exploratory analyses in the prospective arm relative to historical controls for the modified intent-to-treat and intent-to-treat populations. Ratios are adjusted for parity, maternal age, and opioid use. (A) Hazard ratios for gestational age at birth (GAB), neonatal hospital length of stay (NNLOS), and neonatal intensive care unit (NICU) length of stay (NICULOS) reflect the rate of prolonged gestational age at delivery and earlier hospital or NICU discharge in the prospective arm relative to historical controls. (B) Odds ratios for neonatal morbidity and mortality (NMI), preterm birth (PTB), and spontaneous preterm birth (sPTB) at various gestational age cutoffs, reflect the rate of prolonged gestational age at birth or reduction in preterm deliveries relative to historical controls. *Solid vertical lines* indicate the null value of the ratio (1.0).



Supplemental Digital Content

Appendix 1. Neonatal morbidity and mortality index (NMI) scoring and definitions of morbidity.

Composite Neonatal Morbidity Index (NMI) Scale*	
0-to-4 scale integrating neonatal intensive care unit (NICU) length of stay	0 = No events 1 = One event for (RDS, BPD, IVH Grade III or IV, any PVL, proven sepsis, or NEC) or 1-4 days in the NICU, and no neonatal mortality 2 = Two events or from 5-20 days in the NICU, and no neonatal mortality 3 = Three or more events or >20 days in the NICU, and no neonatal mortality 4 = Neonatal mortality
Definitions of Neonatal Morbidity	
Category	**Description**
Intraventricular hemorrhage (IVH)	Determined by cranial ultrasound or computed tomography • Grade I subependymal hemorrhage • Grade II intraventricular hemorrhage, uncomplicated • Grade III intraventricular hemorrhage with ventricular dilatation • Grade IV intraventricular hemorrhage with ventricular dilatation and parenchymal extension
Periventricular leukomalacia (PVL)	Determined by cranial ultrasound • Any PVL • Cystic PVL
Necrotizing enterocolitis (NEC)	• Stage I: Other – Suspect (treatment was observation) • Stage II: Clinical – Definite (treatment was medical) • Stage III: Surgical – Advanced (treatment was surgical)

Respiratory distress syndrome (RDS) or Hyaline membrane disease (HMD)	Requires both diagnosis and oxygen therapy
	• Must include:
	  ○ Oxygen therapy (fraction of inspired oxygen (FiO2) ≥0.40) until infant death or ≥24 hours or continuous positive airway pressure (CPAP) and
	  ○ Clinical diagnosis of RDS or HMD

Retinopathy of prematurity (ROP)	• Stage I: Ophthalmoscopic demarcation line of normal and abnormal vessels • Stage II: Intra-retinal ridge (ridge that rises from the retina as a result of the growth of the abnormal vessels) • Stage III: Ridge with extraretinal fibrovascular proliferation (the ridge grows from the spread of the abnormal vessels and extends into the vitreous) • Stage III+: Stage III and "plus disease", meaning that the blood vessels of the retina have become enlarged and twisted, indicating a worsening of the disease • Stage IV: Partially detached retina • Stage V: Complete retinal detachment
Bronchopulmonary dysplasia (BPD)	• Treatment with >21% oxygen for at least 28 days, or • Oxygen dependence after 36 weeks post-conceptional age

Sepsis	Must include:
	• Blood culture-proven sepsis, and
	• A clinically ill infant with infection defined as:
	  ○ Bacterial sepsis of the newborn
	  ○ Streptococcal sepsis
	  ○ Severe sepsis

Neonatal seizure	Any incident(s) documented as evidence of seizure/epileptic activity by the neonatal staff
Neonatal mortality	Neonatal death within 28 days of delivery

*Adapted from: Hassan SS, Romero R, Vidyadhari D et al.; PREGNANT Trial. Vaginal progesterone reduces the rate of preterm birth in women with a sonographic short cervix: a multicenter, randomized, double-blind, placebo-controlled trial. Ultrasound Obstet Gynecol. 2011 Jul;38(1):18-31. doi: 10.1002/uog.9017. Epub 2011 Jun 15. PMID: 21472815; PMCID: PMC3482512.

Appendix 2. Sensitivity analysis of neonatal morbidity and mortality index severity level probabilities and percent differences between arms for a range of covariate values.

				Predicted probabilities of NMI categories							Percent reduction in risks in Prospective arm relative to Historical arm	
Arm	**Multiparous**	**Age**	**Opioid Use**	**NMI = 0**	**NMI = 1**	**NMI = 2**	**NMI = 3**	**NMI = 4**	**NMI >= 3**	**NMI >= 1**	**NMI >= 3**	**NMI >= 1**
Historical	No	30	No	0.831	0.082	0.055	0.024	0.009	0.033	0.169		
Prospective	No	30	No	0.858	0.069	0.045	0.020	0.007	0.027	0.142	18.47%	16.29%
Historical	Yes	30	No	0.894	0.053	0.034	0.014	0.005	0.020	0.106		
Prospective	Yes	30	No	0.912	0.044	0.028	0.012	0.004	0.016	0.088	18.68%	17.32%
Historical	No	30	Yes	0.652	0.147	0.119	0.060	0.022	0.082	0.348		
Prospective	No	30	Yes	0.698	0.132	0.102	0.050	0.018	0.068	0.302	17.69%	13.24%
Historical	Yes	30	Yes	0.763	0.109	0.078	0.036	0.013	0.050	0.237		
Prospective	Yes	30	Yes	0.799	0.095	0.066	0.030	0.011	0.041	0.201	18.21%	15.16%
Historical	No	20	No	0.842	0.077	0.051	0.023	0.008	0.031	0.158		
Prospective	No	20	No	0.868	0.065	0.042	0.018	0.006	0.025	0.132	18.51%	16.47%
Historical	Yes	20	No	0.901	0.049	0.031	0.013	0.005	0.018	0.099		
Prospective	Yes	20	No	0.919	0.041	0.026	0.011	0.004	0.015	0.081	18.70%	17.44%
Historical	No	20	Yes	0.670	0.141	0.113	0.056	0.021	0.076	0.330		
Prospective	No	20	Yes	0.715	0.127	0.096	0.046	0.017	0.063	0.285	17.79%	13.56%
Historical	Yes	20	Yes	0.777	0.103	0.073	0.034	0.012	0.046	0.223		
Prospective	Yes	20	Yes	0.812	0.090	0.061	0.028	0.010	0.038	0.188	18.27%	15.40%
Historical	No	40	No	0.819	0.087	0.059	0.026	0.009	0.036	0.181		
Prospective	No	40	No	0.848	0.074	0.049	0.022	0.008	0.029	0.152	18.43%	16.10%
Historical	Yes	40	No	0.886	0.057	0.036	0.016	0.005	0.021	0.114		
Prospective	Yes	40	No	0.906	0.047	0.030	0.013	0.004	0.017	0.094	18.65%	17.19%
Historical	No	40	Yes	0.633	0.152	0.126	0.065	0.024	0.089	0.367		
Prospective	No	40	Yes	0.680	0.138	0.109	0.053	0.020	0.073	0.320	17.59%	12.91%
Historical	Yes	40	Yes	0.748	0.115	0.084	0.039	0.014	0.054	0.252		
Prospective	Yes	40	Yes	0.785	0.100	0.070	0.032	0.012	0.044	0.215	18.15%	14.91%

NMI, neonatal morbidity index