HHS Public Access

Author manuscript

N Engl J Med. Author manuscript; available in PMC 2019 August 09.

Published in final edited form as:

N Engl J Med. 2018 August 09; 379(6): 513–523. doi:10.1056/NEJMoa1800566.

Labor Induction versus Expectant Management in Low-Risk Nulliparous Women

William A. Grobman, M.D., **Madeline M. Rice, Ph.D.**, **Uma M. Reddy, M.D.,M.P.H.**, **Alan T.N. Tita, M.D.,Ph.D.**, **Robert M. Silver, M.D.**, **Gail Mallett, R.N.,M.S.,C.C.R.C.**, **Kim Hill, R.N.,B.S.N.**, **Elizabeth A. Thom, Ph.D.**, **Yasser Y. El-Sayed, M.D.**, **Annette Perez-Delboy, M.D.**, **Dwight J. Rouse, M.D.**, **George R. Saade, M.D.**, **Kim A. Boggess, M.D.**, **Suneet P. Chauhan, M.D.**, **Jay D. Iams, M.D.**, **Edward K. Chien, M.D.**, **Brian M. Casey, M.D.**, **Ronald S. Gibbs, M.D.**, **Sindhu K. Srinivas, M.D.,M.S.C.E.**, **Geeta K. Swamy, M.D.**, **Hyagriv N. Simhan, M.D.**, **George A. Macones, M.D.,M.S.C.E.**, and **Eunice Kennedy Shriver National Institute of Child Health and Human Development Maternal–Fetal Medicine Units Network***

Department of Obstetrics and Gynecology, Northwestern University, Chicago (W.A.G., G.M.); University of Alabama at Birmingham, Birmingham (A.T.N.T.); University of Utah Health Sciences Center, Salt Lake City (R.M.S., K.H.); Stanford University, Stanford, CA (Y.Y.E.-S.); Columbia University, New York (A.P.-D.); Brown University, Providence, RI (D.J.R.); University of Texas Medical Branch, Galveston (G.R.S.), University of Texas Health Science Center at Houston, Children's Memorial Hermann Hospital, Houston (S.P.C.), and University of Texas Southwestern Medical Center, Dallas (B.M.C.) — all in Texas; University of North Carolina at Chapel Hill, Chapel Hill (K.A.B.), and Duke University, Durham (G.K.S.) — both in North Carolina; Ohio State University, Columbus (J.D.I.), and MetroHealth Medical Center, Case Western Reserve University, Cleveland (E.K.C.) — both in Ohio; University of Colorado School of Medicine, Anschutz Medical Campus, Aurora (R.S.G.); University of Pennsylvania, Philadelphia (S.K.S.); University of Pittsburgh, Pittsburgh (H.N.S.) — both in Pennsylvania; Washington University, St. Louis (G.A.M.); the George Washington University Biostatistics Center, Washington, DC (M.M.R., E.A.T.); and the Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, MD (U.M.R.).

Abstract

BACKGROUND—The perinatal and maternal consequences of induction of labor at 39 weeks among low-risk nulliparous women are uncertain.

METHODS—In this multicenter trial, we randomly assigned low-risk nulliparous women who were at 38 weeks 0 days to 38 weeks 6 days of gestation to labor induction at 39 weeks 0 days to

*A list of other members of the Eunice Kennedy Shriver National Institute of Child Health and Human Development Maternal–Fetal Medicine Units Network is provided in the Supplementary Appendix, available at NEJM.org.

The authors' affiliations are listed in the Appendix. Address reprint requests to Dr. Grobman at the Department of Obstetrics and Gynecology, Northwestern University, 250 E. Superior St., Suite 05-2175, Chicago, IL 60611, or at w-grobman@northwestern.edu.

Dr. Silver reports receiving consulting fees from Gestavision. No other potential conflict of interest relevant to this article was reported.

39 weeks 4 days or to expectant management. The primary outcome was a composite of perinatal death or severe neonatal complications; the principal secondary outcome was cesarean delivery.

RESULTS—A total of 3062 women were assigned to labor induction, and 3044 were assigned to expectant management. The primary outcome occurred in 4.3% of neonates in the induction group and in 5.4% in the expectant-management group (relative risk, 0.80; 95% confidence interval [CI], 0.64 to 1.00). The frequency of cesarean delivery was significantly lower in the induction group than in the expectant-management group (18.6% vs. 22.2%; relative risk, 0.84; 95% CI, 0.76 to 0.93).

CONCLUSIONS—Induction of labor at 39 weeks in low-risk nulliparous women did not result in a significantly lower frequency of a composite adverse perinatal outcome, but it did result in a significantly lower frequency of cesarean delivery. (Funded by the Eunice Kennedy Shriver National Institute of Child Health and Human Development; ARRIVE ClinicalTrials.gov number, NCT01990612.)

RECOMMEDATION REGAEDING the timing of delivery are founded on a balancing of maternal and perinatal risks. Delivery before 39 weeks 0 days of gestation without medical indication is associated with worse perinatal outcomes than delivery at full term.[1] For women who are at 41 weeks of gestation or later, delivery has been recommended because of increasing perinatal risks.[2] When gestation is between 39 weeks 0 days and 40 weeks 6 days, common practice has been to avoid elective labor induction because of a lack of evidence of peri-natal benefit and concern about a higher frequency of cesarean delivery and other possible adverse maternal outcomes, particularly among nulliparous women.[3]

However, these conclusions were derived largely from observational studies in which labor induction was compared with spontaneous labor.[4–6] Such a comparison provides little insight into clinical management, because spontaneous labor is not a certain alternative to labor induction. Most observational studies that have used the clinically relevant comparator of expectant management have not shown a higher risk of adverse outcomes with labor induction; instead, some of these studies have shown that induction of labor resulted in a lower frequency of cesarean delivery and more favorable perinatal outcomes than expectant management.[7–11]

A previous randomized trial conducted in the United Kingdom compared labor induction at 39 weeks of gestation with expectant management among 619 women who were 35 years of age or older and who had no other indication for delivery at 39 weeks of gestation.[12] The frequency of cesarean delivery was similar in the two groups (relative risk, 0.99; 95% confidence interval [CI], 0.87 to 1.14), although several aspects of the trial, including a rate of operative vaginal delivery (i.e., vaginal delivery with the use of forceps or vacuum) of more than 30%, called into question the external validity of these results for the United States. The authors of that trial encouraged replication of their findings in other populations and the performance of a trial with a sample size sufficient “to test the effects of induction on uncommon adverse neonatal outcomes.” The ARRIVE trial (A Randomized Trial of Induction Versus Expectant Management) was designed to test the hypothesis that elective induction of labor at 39 weeks would result in a lower risk of a composite outcome of

perinatal death or severe neonatal complications than expectant management among low-risk nulliparous women.

METHODS

TRIAL OVERSIGHT

We conducted this multicenter, randomized, controlled, parallel-group, unmasked trial at 41 hospitals participating in the Maternal–Fetal Medicine Units Network of the Eunice Kennedy Shriver National Institute of Child Health and Human Development. The protocol (available with the full text of this article at NEJM.org) was approved by the institutional review board at each hospital before participant enrollment. Written informed consent was obtained from all participants before randomization. An independent data and safety monitoring committee monitored the trial. The authors vouch for the accuracy and completeness of the data and for the fidelity of the trial to the protocol.

SCREENING AND RECRUITMENT

Low-risk nulliparous women who were at 34 weeks 0 days to 38 weeks 6 days of gestation with a live singleton fetus that was in a vertex presentation, who had no contraindication to vaginal delivery, and who had no cesarean delivery planned were screened for eligibility. Low risk was defined as the absence of any condition considered to be a maternal or fetal indication for delivery before 40 weeks 5 days (e.g., hypertensive disorders of pregnancy or suspected fetal-growth restriction). Reliable information on the length of gestation was also a criterion for enrollment; information was considered to be reliable if the woman was certain of the date of her last menstrual period and that date was consistent with results of ultrasonography performed before 21 weeks 0 days or if the date of the last menstrual period was uncertain but results were available from ultrasonography performed before 14 weeks 0 days. Full eligibility criteria are provided in the Supplementary Appendix, available at NEJM.org.

RANDOMIZATION AND MANAGEMENT STRATEGY

Women who consented to participate were assessed again between 38 weeks 0 days and 38 weeks 6 days of gestation to ensure that they did not have new indications for delivery that would make them ineligible for the trial. Women who were in labor or had premature rupture of membranes or vaginal bleeding at this time were considered to be ineligible. Women who met the inclusion criteria were randomly assigned in a 1:1 ratio to either labor induction or expectant management. The randomization sequence, prepared by an independent data coordinating center, used the simple urn method, with stratification according to clinical site.[13] The cervix was examined before labor, from 72 hours before to 24 hours after randomization, to assess dilation, effacement, and station of the fetus to determine a modified Bishop score (scores range from 0 to 12, with lower scores associated with a higher chance of cesarean delivery) (see the Supplementary Appendix).[14]

Women in the induction group were assigned to undergo induction of labor at 39 weeks 0 days to 39 weeks 4 days. Women in the expectant-management group were asked to forego elective delivery before 40 weeks 5 days and to have delivery initiated no later than 42

weeks 2 days. A specific induction protocol was not mandated for women who underwent induction in either group. Other protocol guidelines are provided in the Supplementary Appendix.

Trained and certified research staff members abstracted information from medical records, including demographic information, medical history, and outcome data. Participants were followed up with an interview performed by research personnel immediately post partum. During this interview, women were asked to rate their labor pain on a 10-point Likert scale (with higher scores indicating greater pain)[15] and to rate their experiences on the Labor Agentry Scale,[16] which was designed to assess expectations and experiences of personal control during childbirth (scores range from 29 to 203, with higher scores indicating greater perceived control during childbirth). The score on the Labor Agentry Scale was also assessed in a second interview performed by research personnel 4 to 8 weeks after delivery.

TRIAL OUTCOMES

The primary outcome was a composite of peri-natal death or severe neonatal complications and consisted of one or more of the following during the antepartum or intrapartum period or during the delivery hospitalization: perinatal death, the need for respiratory support within 72 hours after birth, Apgar score of 3 or less at 5 minutes, hypoxic–ischemic encephalopathy,[17] seizure, infection (confirmed sepsis or pneumonia), meconium aspiration syndrome, birth trauma (bone fracture, neurologic injury, or retinal hemorrhage), intracranial or subgaleal hemorrhage, or hypo-tension requiring vasopressor support. The principal prespecified maternal outcome (the main secondary outcome) was cesarean delivery.

Prespecified subgroups for the primary perinatal outcome and for the secondary outcome of cesarean delivery were maternal race or ethnic group as reported by the participant (white, black, Asian, Hispanic, other, unknown, or more than one race), age of 35 years or older versus younger than 35 years, body-mass index (the weight in kilograms divided by the square of the height in meters) of 30 or more versus less than 30, and a modified Bishop score at the time of randomization of less than 5 versus 5 or higher. In addition, although it was not a baseline variable, the specialty of the admitting provider (obstetrics-gynecology, maternal-fetal medicine, family practice, or midwifery) was prespecified for the subgroup analyses.

Neonatal secondary outcomes included birth weight, duration of respiratory support, cephalohematoma, shoulder dystocia, transfusion of blood products, hyperbilirubinemia requiring photo-therapy or exchange transfusion, hypoglycemia requiring intravenous therapy, admission to the neonatal intermediate or intensive care unit, and length of hospitalization. In addition to cesarean delivery, other maternal secondary outcomes included hypertensive disorders of pregnancy (gestational hypertension or preeclampsia), indication for cesarean delivery, operative vaginal delivery, indication for operative vaginal delivery, uterine incisional extensions during cesarean delivery, chorioamnionitis, third-degree or fourth-degree perineal laceration, postpartum hemorrhage, postpartum infection, venous thromboembolism, number of hours in the labor and delivery unit, length of postpartum hospital stay, admission to the intensive care unit, and maternal death. Definitions of secondary outcomes are provided in the Supplementary Appendix.

Records of all infants who met the primary perinatal outcome were reviewed centrally to verify that the primary outcome had occurred. Records of infants in whom the primary outcome did not occur but that suggested (on the basis of a delivery hospitalization of 7 or more days or discharge to a long-term care facility) that clinically significant perinatal complications may have occurred were reviewed centrally as well. Reviewers were unaware of the trial-group assignments.

STATISTICAL ANALYSIS

The expected rate of the primary perinatal outcome in the expectant-management group was estimated to be 3.5%.[18] We calculated that enrollment of 6000 women would provide a power of at least 85% to detect a 40% lower rate of the primary outcome in the induction group than in the expectant-management group, at a two-sided type I error rate of 5%. This power analysis incorporated the assumption that for 7.5% of the women, management would not be consistent with the protocol of the assigned strategy.

Analyses were performed according to the intention-to-treat principle. We compared continuous variables using the Wilcoxon signed-rank test and categorical variables using the chi-square and Fisher's exact tests. A multinomial outcome was compared with the use of multi-nomial logistic regression. Time variables measured in days were categorized and compared with the Cochran–Armitage trend test. We used a group sequential method to control the type I error with the Lan–DeMets characterization of the O'Brien–Fleming boundary. One interim analysis was performed; in the final analysis of the primary outcome, a two-tailed P value of less than 0.046 was considered to indicate statistical significance. Because the adjustment is minimal, we report the 95% confidence interval for the relative risk. Our statistical analysis plan did not call for adjustment of P values to control for multiple comparisons of the results for the individual components of the primary outcome; therefore, these are reported as point estimates and 95% confidence intervals. For the secondary outcomes, the level of significance was adjusted post hoc for multiple comparisons with the false discovery rate method.[19] No method of imputation of missing data was used, although sensitivity analyses were performed in which data from participants who withdrew consent or were lost to follow-up were handled in various ways. To determine whether there was a differential effect of labor induction on the primary perinatal outcome and on the secondary outcome of cesarean delivery within the prespecified subgroups, we performed the Breslow–Day interaction test in which a P value of less than 0.05 was considered to indicate statistical significance. The statistical analysis plan is provided in the protocol, available at NEJM.org.

RESULTS

CHARACTERISTICS OF THE PARTICIPANTS

From March 2014 through August 2017, a total of 50,581 women underwent screening for eligibility. Of the 22,533 eligible women, 6106 (27%) provided written informed consent and underwent randomization: 3062 were assigned to the induction group, and 3044 to the expectant-management group (Fig. 1). At the time of randomization, 63% of the participants had an unfavorable modified Bishop score (i.e., a score <5). The two groups were similar at

baseline, except that fewer women in the induction group than in the expectant-management group had had a previous pregnancy loss (22.8% vs. 25.6%, P = 0.01) (Table 1). The obstetrical provider at the time of admission for delivery was a physician for 94% of women and a midwife for 6%.

ADHERENCE

Three women in the induction group and 7 in the expectant-management group were lost to follow-up or withdrew consent. In the case of 184 women (6.0%) in the induction group and 140 (4.6%) in the expectant-management group, the management was not consistent with the protocol of the assigned strategy (details are provided in the Supplementary Appendix). Women in the induction group had a shorter median time from randomization to delivery than women in the expectant-management group (7 days [interquar-tile range, 5 to 9] vs. 12 days [interquartile range, 7 to 16], P<0.001); in addition, women in the induction group underwent delivery at a significantly earlier median gestational age (39.3 weeks [interquartile range, 39.1 to 39.6] vs. 40.0 weeks [interquartile range, 39.3 to 40.7], P<0.001) and had neonates with significantly lower median birth weights (3300 g [interquartile range, 3040 to 3565] vs. 3380 g [interquartile range, 3110 to 3650], P<0.001).

PRIMARY OUTCOME AND OTHER PERINATAL OUTCOMES

The primary perinatal outcome occurred in 4.3% of the neonates in the induction group and in 5.4% in the expectant-management group (relative risk, 0.80; 95% CI, 0.64 to 1.00; P = 0.049 [P<0.046 indicated statistical significance for the primary perinatal outcome]) (Table 2). This finding did not change after adjustment for previous pregnancy loss and was materially unchanged in the sensitivity analyses. Neonates in the induction group also had a shorter duration of respiratory support and of total hospital stay. Other secondary perinatal outcomes were similar in the two groups (see the Supplementary Appendix).

MATERNAL OUTCOMES

The percentage of women who underwent cesarean delivery was significantly lower in the induction group than in the expectant-management group (18.6% vs. 22.2%; relative risk, 0.84; 95% CI, 0.76 to 0.93; P<0.001) (Table 3). This finding did not change materially after adjustment for previous pregnancy loss. Women assigned to induction of labor were also significantly less likely than women assigned to expectant management to have hypertensive disorders of pregnancy (9.1% vs. 14.1%; relative risk, 0.64; 95% CI, 0.56 to 0.74; P<0.001) and to have extensions of the uterine incision during cesarean delivery; in addition, women in the induction group reported less pain (i.e., had lower scores on the 10-point Likert scale) and more perceived control during childbirth (i.e., had higher scores on the Labor Agentry Scale). Although differences in scores were statistically significant, they were relatively small. Women in the induction group spent more time in the labor and delivery unit, but the length of their postpartum hospital stay was shorter (Table 3). Other secondary maternal health outcomes were similar in the two groups (see the Supplementary Appendix).

SUBGROUP ANALYSES

Prespecified baseline subgroup analyses of the primary perinatal outcome and of the secondary outcome of cesarean delivery showed no significant differences in results according to race or ethnic group, maternal age, body-mass index, or modified Bishop score (all P>0.05 by the Breslow– Day test for homogeneity) (Fig. 2). Subgroup analysis also revealed no significant between-group difference in the two outcomes according to type of admitting provider.

DISCUSSION

In this randomized trial involving low-risk nulliparous women, we did not find a significant difference in the frequency of the primary outcome (a composite of adverse perinatal outcomes) between women randomly assigned to labor induction at 39 weeks of gestation and women assigned to expectant management. Nevertheless, the relative risk was 20% lower in the induction group than in the expectant-management group, and the corresponding 95% confidence interval suggests that labor induction is probably not associated with a higher risk of adverse perinatal outcomes than expectant management, and it may be associated with as much as a 36% lower risk than expectant management. Labor induction also resulted in a significantly lower frequency of cesarean delivery and hypertensive disorders of pregnancy than expectant management, even after post hoc adjustment for multiplicity. Our data suggest that 1 cesarean delivery may be avoided for every 28 deliveries among low-risk nulliparous women who plan to undergo elective induction of labor at 39 weeks.

These findings contradict the conclusions of multiple observational studies that have suggested that labor induction is associated with an increased risk of adverse maternal and perinatal outcomes.[4–6] These studies, however, compared women who underwent labor induction with those who had spontaneous labor, which is not a comparison that is useful to guide clinical decision making. Conversely, our findings are consistent with observational studies,[7–11,20–23] as well as the randomized trial conducted by Walker et al.,[12] in which women undergoing labor induction were compared with women undergoing the actual clinical alternative of expectant management.

We found no significant difference in the magnitude of effect with respect to the primary perinatal outcome or cesarean delivery according to whether a woman had an unfavorable modified Bishop score at randomization. This finding may seem unexpected, given the consistent evidence that women with an unfavorable Bishop score have a higher chance of cesarean delivery when labor is induced than women with a favorable score.[3] As shown by the frequency of cesarean delivery among women with an unfavorable as opposed to a favorable baseline modified Bishop score (i.e., a score ≥5), this relationship holds true in our trial. Yet, because women with an unfavorable score at baseline also had a higher chance of cesarean delivery than women with a favorable score when they followed the expectant-management strategy, labor induction in women with an unfavorable score still resulted in fewer cesarean deliveries than expectant management.

This trial is larger than previous randomized trials that compared labor induction with expectant management in low-risk women, and as such it had the ability to detect differences that previous trials may not have discerned. Eligibility criteria ensured that only women with reliable information on length of gestation were included, and both women with favorable modified Bishop scores at baseline and those with unfavorable scores were enrolled.

Limitations of the trial should be noted. First, because masking was not feasible, ascertainment bias is possible. Second, despite its size, the trial was not powered to detect differences in infrequent outcomes, and most individual adverse perinatal outcomes were relatively uncommon. Third, it is unclear whether results are broadly generalizable; however, the inclusion of both university and community hospitals throughout the United States and of a variety of types of obstetrical providers, as well as the absence of a single protocol for induction or labor management, suggests that results are probably general-izable to similar centers. Finally, the cost-effectiveness of labor induction in low-risk nulliparous women at 39 weeks will need to be evaluated in further analyses.

In summary, we found that elective labor induction at 39 weeks of gestation did not result in a greater frequency of perinatal adverse outcomes than expectant management and resulted in fewer instances of cesarean delivery. These results suggest that policies aimed at the avoidance of elective labor induction among low-risk nulliparous women at 39 weeks of gestation are unlikely to reduce the rate of cesarean delivery on a population level; the trial provides information that can be incorporated into discussions that rely on principles of shared decision making.[24–27]

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank Lindsay Doherty, M.S., for managing the data and protocol and for performing statistical analysis.

References

1. Parikh LI, Reddy UM, Männistö T, et al. Neonatal outcomes in early term birth. Am J Obstet Gynecol 2014; 211(3): 265.e1–265.e11. [PubMed: 24631438]
2. American College of Obstetricians and Gynecologists. Practice bulletin no. 146: management of late-term and postterm pregnancies. Obstet Gynecol 2014; 124: 390–6. [PubMed: 25050770]
3. ACOG Committee on Practice Bulletins — Obstetrics. ACOG practice bulletin no. 107: induction of labor. Obstet Gynecol 2009; 114: 386–97. [PubMed: 19623003]
4. Vardo JH, Thornburg LL, Glantz JC. Maternal and neonatal morbidity among nulliparous women undergoing elective induction of labor. J Reprod Med 2011; 56: 25–30. [PubMed: 21366123]
5. Dunne C, Da Silva O, Schmidt G, Natale R Outcomes of elective labour induction and elective caesarean section in low-risk pregnancies between 37 and 41 weeks' gestation. J Obstet Gynaecol Can 2009; 31: 1124–30. [PubMed: 20085677]

6. Guerra GV, Cecatti JG, Souza JP, et al. Elective induction versus spontaneous labour in Latin America. Bull World Health Organ 2011; 89: 657–65. [PubMed: 21897486]
7. Osmundson S, Ou-Yang RJ, Grobman WA. Elective induction compared with expectant management in nulliparous women with an unfavorable cervix. Obstet Gynecol 2011; 117: 583–7. [PubMed: 21343761]
8. Gibson KS, Waters TP, Bailit JL. Maternal and neonatal outcomes in electively induced low-risk term pregnancies. Am J Obstet Gynecol 2014; 211(3): 249.e1–249.e16. [PubMed: 24631440]
9. Stock SJ, Ferguson E, Duffy A, Ford I, Chalmers J, Norman JE. Outcomes of elective induction of labour compared with expectant management: population based study. BMJ 2012; 344: e2838. [PubMed: 22577197]
10. Cheng YW, Kaimal AJ, Snowden JM, Nicholson JM, Caughey AB. Induction of labor compared to expectant management in low-risk women and associated perinatal outcomes. Am J Obstet Gynecol 2012; 207(6): 502.e1–502.e8. [PubMed: 23063017]
11. Darney BG, Snowden JM, Cheng YW, et al. Elective induction of labor at term compared with expectant management: maternal and neonatal outcomes. Obstet Gynecol 2013; 122: 761–9. [PubMed: 24084532]
12. Walker KF, Bugg GJ, Macpherson M, et al. Randomized trial of labor induction in women 35 years of age or older. N Engl J Med 2016; 374: 813–22. [PubMed: 26962902]
13. Rosenberger W, Lachin JM. Randomization in clinical trials New York: Wiley, 2002.
14. Lange AP, Secher NJ, Westergaard JG, Skovgård I. Prelabor evaluation of inducibility. Obstet Gynecol 1982; 60: 137–47. [PubMed: 7155472]
15. Ludington E, Dexter F. Statistical analysis of total labor pain using the visual analog scale and application to studies of analgesic effectiveness during childbirth. Anesth Analg 1998; 87: 723–7. [PubMed: 9728862]
16. Hodnett ED, Simmons-Tropea DA. The Labour Agentry Scale: psychometric properties of an instrument measuring control during childbirth. Res Nurs Health 1987; 10(5): 301–10. [PubMed: 3671777]
17. Shankaran S, Laptook AR, Ehrenkranz RA, et al. Whole-body hypothermia for neonates with hypoxic-ischemic encephalopathy. N Engl J Med 2005; 353: 1574–84. [PubMed: 16221780]
18. Bailit JL, Grobman WA, Rice MM, et al. Risk-adjusted models for adverse obstetric outcomes and variation in risk-adjusted outcomes across hospitals. Am J Obstet Gynecol 2013; 209(5): 446.e1–446.e30. [PubMed: 23891630]
19. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc [B] 1995; 57: 289–300.
20. Caughey AB, Sundaram V, Kaimal AJ, et al. Maternal and neonatal outcomes of elective induction of labor. Evid Rep Technol Assess (Full Rep) 2009; 176: 1–257.
21. Caughey AB, Sundaram V, Kaimal AJ, et al. Systematic review: elective induction of labor versus expectant management of pregnancy. Ann Intern Med 2009; 151: 252–263, W53–W63. [PubMed: 19687492]
22. Kawakita T, Iqbal SN, Huang CC, Reddy UM. Nonmedically indicated induction in morbidly obese women is not associated with an increased risk of cesarean delivery. Am J Obstet Gynecol 2017; 217(4): 451.e1–451.e8. [PubMed: 28578171]
23. Gibson KS, Waters TP, Bailit JL. A risk of waiting: the weekly incidence of hyper-tensive disorders and associated maternal and neonatal morbidity in low-risk term pregnancies. Am J Obstet Gynecol 2016; 214(3): 389.e1–389.e12. [PubMed: 26440691]
24. Sparks TN, Yeaton-Massey A, Granados JM, Handler SJ, Meyer M, Caughey AB. How do maternal views of delivery outcomes vary by demographics and preferred mode of delivery? Am J Perinatol 2015; 32: 741–6. [PubMed: 25519201]
25. Schwarz C, Gross MM, Heusser P, Berger B. Women's perceptions of induction of labour outcomes: results of an online-survey in Germany. Midwifery 2016; 35: 3–10. [PubMed: 27060393]
26. Moore JE, Low LK, Titler MG, Dalton VK, Sampselle CM. Moving toward patient-centered care: women's decisions, perceptions, and experiences of the induction of labor process. Birth 2014; 41: 138–46. [PubMed: 24702312]

27. Yee LM, Kaimal AJ, Houston KA, et al. Mode of delivery preferences in a diverse population of pregnant women. Am J Obstet Gynecol 2015; 212(3): 377.e1–377.e24. [PubMed: 25446662]

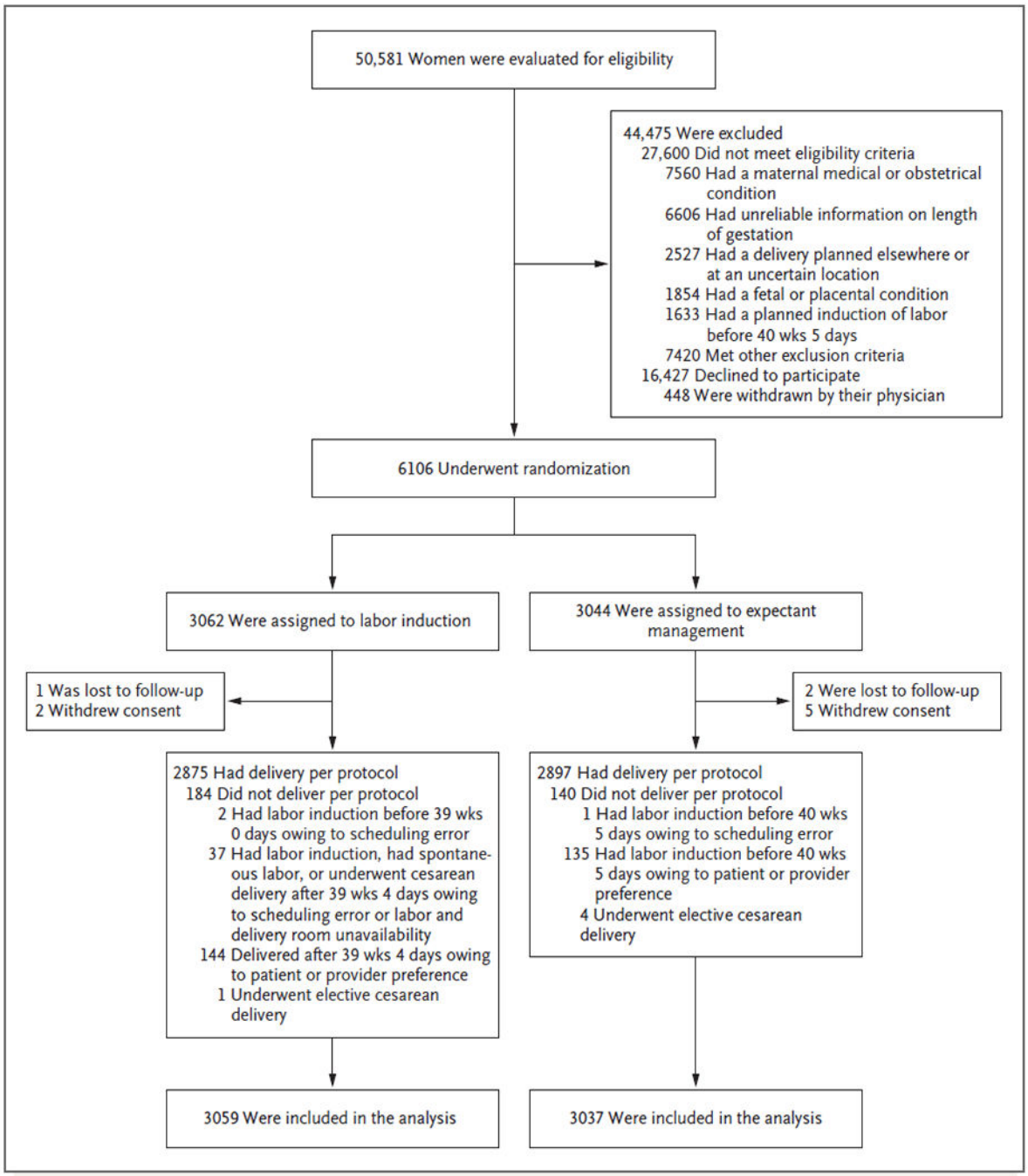

Figure 1. Eligibility, Randomization, Delivery, and Assessment.

Per-protocol delivery in the induction group was defined as electively induced labor from 39 weeks 0 days to 39 weeks 4 days or spontaneous labor or medically indicated delivery on or before 39 weeks 4 days (this also included delivery delayed past 39 weeks 4 days because of a new medical indication that had developed). Per-protocol delivery in the expectant-management group was defined as induction from 40 weeks 5 days to 42 weeks 2 days or spontaneous or medically indicated delivery on or before 42 weeks 2 days.

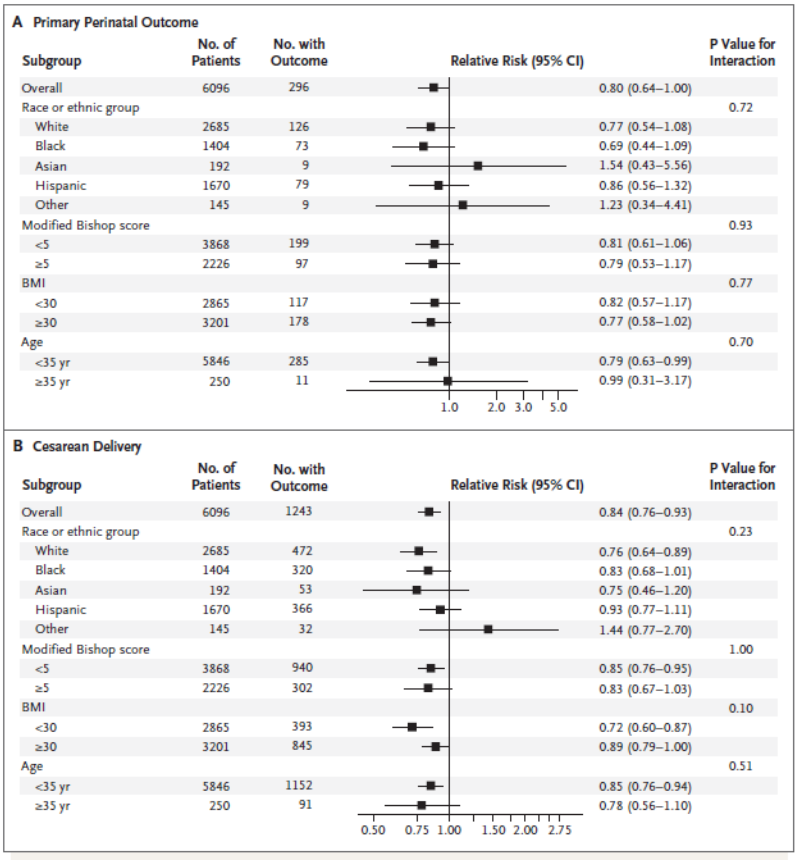

Figure 2. Prespecified Subgroup Analyses According to Maternal Baseline Variables.

The primary outcome was a composite of perinatal death or severe neonatal complications and consisted of one or more of the following during the antepartum or intrapartum period or during the delivery hospitalization: perinatal death, the need for respiratory support within the first 72 hours after birth, Apgar score of 3 or less at 5 minutes, hypoxic–ischemic encephalopathy, seizure, infection (confirmed sepsis or pneumonia), meconium aspiration syndrome, birth trauma (bone fracture, neurologic injury, or retinal hemorrhage), intracranial or subgaleal hemorrhage, or hypotension requiring vasopressor support. Race was reported by the participant; “other” race or ethnic group includes other, unknown, or more than one race or ethnic group. Modified Bishop scores range from 0 to 12, with lower scores associated with a higher chance of cesarean delivery. The body-mass index (BMI) is the weight in kilograms divided by the square of the height in meters.

Table 1.

Maternal Characteristics at Baseline.*

Characteristic	Induction Group (N = 3062)	Expectant-Management Group (N = 3044)
Age — yr		
Median	24	23
Interquartile range	21–28	20–28
Age ≥35 yr — no. (%)	114 (3.7)	136 (4.5)
Race or ethnic group — no. (%)†		
White	1329 (43.4)	1359 (44.6)
Black	707 (23.1)	699 (23.0)
Asian	87 (2.8)	106 (3.5)
Hispanic	866 (28.3)	808 (26.5)
Other, unknown, or more than one race	73 (2.4)	72 (2.4)
Married or living with a partner — no. (%)	1814 (59.2)	1798 (59.1)
Employment status — no./total no. (%)‡		
Employed full time	1226/3053 (40.2)	1209/3036 (39.8)
Employed part time	341/3053 (11.2)	353/3036 (11.6)
Not employed	1486/3053 (48.7)	1474/3036 (48.6)
Had private insurance for prenatal care — no./total no. (%)§	1404/3061 (45.9)	1335/3044 (43.9)
History of pregnancy loss — no. (%)		
No previous pregnancy loss	2364 (77.2)	2266 (74.4)
Previous pregnancy loss	698 (22.8)	778 (25.6)
Before 13 wk of gestation only	637 (20.8)	698 (22.9)
At 13–19 wk of gestation only	23 (0.8)	40 (1.3)
Both before 13 wk and at 13–19 wk of gestation	14 (0.5)	17 (0.6)
Ectopic or molar pregnancy only	24 (0.8)	21 (0.7)
Uncertain time of pregnancy loss	0	2 (0.1)
Length of gestation at randomization — wk		
Median	38.3	38.3
Interquartile range	38.0–38.6	38.0–38.6
Method of conception — no. (%)		
In vitro fertilization	56 (1.8)	47 (1.5)
Ovulation induction or artificial insemination	30 (1.0)	24 (0.8)
Spontaneous	2976 (97.2)	2973 (97.7)
Smoked cigarettes — no. (%)	224 (7.3)	242 (8.0)
Drank alcohol — no./total no. (%)¶	133/3062 (4.3)	107/3043 (3.5)
BMI at randomization‖		
Median	30.5	30.3
Interquartile range	27.3–34.6	27.3–35.0
BMI ≥30 — no./total no. (%)‖	1632/3049 (53.5)	1575/3027 (52.0)
Modified Bishop score at randomization**		

Characteristic	Induction Group (N = 3062)	Expectant-Management Group (N = 3044)
Median	4	4
Interquartile range	2–5	2–5
Score <5 — no./total no. (%)**	1919/3062 (62.7)	1954/3042 (64.2)

*There were no significant differences between the groups except for previous pregnancy loss, which was less common in the induction group (P = 0.01). Percentages may not total 100 because of rounding.

†Race or ethnic group was reported by the participant.

‡Data are missing for 17 women (9 in the induction group and 8 in the expectant-management group).

§Data are missing for 1 woman in the induction group.

¶Data are missing for 1 woman in the expectant-management group.

‖The body-mass index (BMI) is the weight in kilograms divided by the square of the height in meters. Data are missing for 30 women (13 in the induction group and 17 in the expectant-management group).

**Modified Bishop scores range from 0 to 12, with lower scores associated with a higher chance of cesarean delivery. Data are missing for 2 women in the expectant-management group.

Table 2.

Primary Perinatal Outcome and Components.*

Outcome	Induction Group (N = 3059)	Expectant-Management Group (N = 3037)	Relative Risk (95% CI)†	P Value‡
	no.(%)			
Primary composite outcome	132 (4.3)	164 (5.4)	0.80 (0.64–1.00)	0.049
Perinatal death	2 (0.1)	3 (0.1)	0.66 (0.12–3.33)	
Respiratory support	91 (3.0)	127 (4.2)	0.71 (0.55–0.93)	
Apgar score ≤3 at 5 min	12 (0.4)	18 (0.6)	0.66 (0.32–1.37)	
Hypoxic-ischemic encephalopathy	14 (0.5)	20 (0.7)	0.70 (0.35–1.37)	
Seizure	11 (0.4)	4 (0.1)	2.74 (0.91–8.12)	
Infection	9 (0.3)	12 (0.4)	0.74 (0.31–1.76)	
Meconium aspiration syndrome	17 (0.6)	26 (0.9)	0.65 (0.35–1.19)	
Birth trauma	14 (0.5)	18 (0.6)	0.77 (0.38–1.55)	
Intracranial or subgaleal hemorrhage	9 (0.3)	7 (0.2)	1.28 (0.48–3.42)	
Hypotension requiring vasopressor support	2 (0.1)	5 (0.2)	0.40 (0.06–1.79)	

*Details regarding the components of the primary perinatal outcome are provided in the Supplementary Appendix.

†Exact confidence intervals are provided for rare outcomes. The widths of the confidence intervals for components of the primary outcome have not been adjusted for multiplicity, so they should not be used to infer definitive effects of the management strategies.

‡We used a group sequential method to control the type I error with the Lan–DeMets characterization of the O'Brien–Fleming boundary. One interim analysis was performed; in the final analysis of the primary outcome, a two-tailed P value of less than 0.046 was considered to indicate statistical significance. Since the adjustment is minimal, we report the 95% confidence interval for the relative risk.

Table 3.

Secondary Outcomes.*

Outcome	Induction Group (N = 3059)	Expectant-Management Group (N = 3037)	Relative Risk (95% CI)	P Value
Neonatal				
Transfusion of blood products — no. (%)	4 (0.1)	5 (0.2)	0.79 (0.20–2.74)	0.75
Hyperbilirubinemia — no. (%)†	145 (4.7)	142 (4.7)	1.01 (0.81–1.27)	0.91
Hypoglycemia — no. (%)	37 (1.2)	35 (1.2)	1.05 (0.66–1.66)	0.84
Admission to neonatal intermediate or intensive care unit — no. (%)	358 (11.7)	394 (13.0)	0.90 (0.79–1.03)	0.13
Maternal				
Cesarean delivery — no. (%)	569 (18.6)	674 (22.2)	0.84 (0.76–0.93)	<0.001‡
Operative vaginal delivery — no. (%)	222 (7.3)	258 (8.5)	0.85 (0.72–1.01)	0.07
Hypertensive disorder of pregnancy — no. (%)	277 (9.1)	427 (14.1)	0.64 (0.56–0.74)	<0.001‡
Chorioamnionitis — no. (%)	407 (13.3)	429 (14.1)	0.94 (0.83–1.07)	0.35
Third-degree or fourth-degree perineal laceration — no. (%)	103 (3.4)	89 (2.9)	1.15 (0.87–1.52)	0.33
Postpartum hemorrhage — no. (%)	142 (4.6)	137 (4.5)	1.03 (0.82–1.29)	0.81
Postpartum infection — no. (%)	50 (1.6)	65 (2.1)	0.76 (0.53–1.10)	0.15
Admission to ICU — no. (%)	4 (0.1)	8 (0.3)	0.50 (0.13–1.55)	0.26
Death — no. (%)	0	0	NA	NA
Median duration of stay in labor and delivery unit (IQR) — hr§	20 (13–28)	14 (9–20)		<0.001‡
Postpartum hospital stay — no. (%)				0.01‡¶
<2 days	322 (10.5)	317 (10.4)		
2 days	2191 (71.6)	2084 (68.6)		
3 days	399 (13.0)	452 (14.9)		
4 days	130 (4.2)	166 (5.5)		
>4 days	17 (0.6)	18 (0.6)		
Median scores on Labor Agentry Scale (IQR)‖				
At 6–96 hr after delivery	168 (148–183)	164 (143–181)		<0.001‡
At 4–8 wk after delivery	176 (157–189)	174 (154–188)		0.01‡
Median labor pain scores (IQR)**				
Worst score	8 (7–10)	9 (8–10)		<0.001‡
Overall score	7 (5–8)	7 (5–9)		<0.001‡

*Additional secondary outcomes are provided in the Supplementary Appendix. Exact confidence intervals and P values are provided for rare outcomes. The P values and 95% confidence intervals presented have not been adjusted for multiple comparisons of the secondary out–comes. ICU denotes intensive care unit, IQR interquartile range, and NA not applicable.

†Data are missing for 4 women (1 in the induction group and 3 in the expectant-management group).

‡The P value remained significant after controlling for multiple comparisons with the false discovery rate method.

§The totals exclude 7 women who delivered before admission to the labor and delivery unit. Data are missing for 2 women (1 in each group).

[¶] The variables were compared with the Cochran–Armitage trend test.

[‖] Scores on the Labor Agentry Scale range from 29 to 203, with higher scores indicating greater perceived control during childbirth; included are women who had spontaneous labor, labor that started spontaneously but then was augmented, or induced labor. Data for 6 to 96 hours after delivery are missing for 288 women (127 in the induction group and 161 in the expectant-management group); data for 4 to 8 weeks after delivery are missing for 736 women (349 in the induction group and 387 in the expectant-management group).

[**] Labor pain was scored according to a 10-point Likert scale, with higher scores indicating greater pain; included are women who had spontaneous labor, labor that started spontaneously but then was augmented, or induced labor. Data on worst score are missing for 274 women (110 in the induction group and 164 in the expectant-management group); data on overall score are missing for 275 women (110 in the induction group and 165 in the expectant-management group).