Published in final edited form as:

Semin Perinatol. 2016 August ; 40(5): 328–334. doi:10.1053/j.semperi.2016.03.011.

What We Have Learned About the Design of Randomized Trials in Pregnancy

Elizabeth A. Thom, PhD¹, Madeline Murguia Rice, PhD¹, George R. Saade, MD², and Uma M. Reddy, MD, MPH³ for the Eunice Kennedy Shriver National Institute of Child Health and Human Development Maternal-Fetal Medicine Units Network

¹Biostatistics Center, George Washington University, Washington DC

²Department of Obstetrics and Gynecology, University of Texas Medical Branch, Galveston, TX

³Eunice Kennedy Shriver National Institute of Child Health and Human Development, Bethesda, MD

Abstract

For nearly 30 years the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) Maternal-Fetal Medicine Units (MFMU) Network has been conducting randomized trials in pregnant women, many of which have changed clinical practice. Since 1986, the MFMU Network has conducted 29 randomized trials, of which the 17 trials started or completed since 2003 are described here. Study design choices are described including decisions regarding the fundamental questions to be answered and the rationale behind choices of primary and secondary outcomes. Some of the potential pitfalls, particularly relating to bias, that can affect the interpretation of trial results are described along with the mechanisms that the Network has used to avoid or minimize them.

Background

In the early days of randomized trials, women, let alone pregnant women, were rarely included. For medical conditions that affected women as well as men it was assumed that the treatment effect of a medication would be similar in women to those in men. As false an assumption as this might have been, it was even more unlikely to apply to pregnant women whose physiological state is quite different from non-pregnant women.¹ At the same time, the traditional reluctance to include pregnant women in randomized trials out of concern for maternal and fetal safety led to use of interventions and medications with unknown risk and the paradoxical situation of exposing more maternal-fetal dyads to potentially harmful interventions than if they had been enrolled in randomized trials. In his essay “Discovering the need for randomized controlled trials in obstetrics: a personal odyssey”, Grimes describes how clinical interventions based only upon opinion or dogma, without solid evidence of benefit, permeated the practice of obstetrics.² The use of tocolysis and electronic

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

fetal heart rate monitoring could be considered as examples. In 1979, Cochrane, the British epidemiologist and early champion of randomized trials, ranked the various medical specialties by the extent to which practices were based on valid evidence of effectiveness. Obstetrics was ranked easily in last place. Not only were randomized trials lacking for obstetrical interventions and management, but the strategies to treat pregnant women with pre-existing diseases were also lacking.³

In the past 40 years there has been a ‘sea change’ in regulations regarding inclusion of women in trials, including the establishment of the Office of Research for Women’s Health. Partly as a result of Cochrane’s observation, the urgent need for randomized trials specifically in obstetrics and maternal-fetal medicine and large enough to give reliable answers, was recognized.⁴ In 1986, the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development (NICHD) created the MFMU Network.⁵ The goal of the Network is to provide the rationale for evidence-based obstetric practice with priority given to randomized trials. Similar initiatives were started in the UK and Canada. To date, the Network has worked on 29 randomized trials: 25 completed, three currently recruiting, and one in the process of implementation. In 2003, this journal published an issue on highlights from the MFMU Network, including lessons learned from the experience of conducting the first 12 trials.⁶ In this article we describe some of the study design challenges particularly relevant to trials in pregnancy that we have faced in the next 17 randomized trials and how we have attempted to overcome them.

Choice of Primary and Secondary Outcomes

For any randomized trial, a primary outcome should be specified (or more than one, as long as the statistical effect of multiplicity is adequately addressed). A good primary outcome is one that is clinically relevant and compelling, sensitive (i.e. likely to be responsive to the intervention), measured precisely and reliably, and measurable in all participants. Choosing a primary outcome for randomized trials in pregnancy can be complicated by the fact that the intervention to be evaluated is administered to one individual – the mother – but is frequently intended to benefit the other(s) – the fetus or fetuses. The balance of safety and effectiveness between the mother and the baby creates a challenge.

Of the 17 trials since 2003 shown in Table 1, 12 (Trials 1, 2, 6–12, 14, 16, 17 in Table 1) were focused primarily on the fetus or infant. In each case, the intervention is at best inconvenient and at worst risky for the women. However, the primary question of the trial is really about the baby. Six of the 12 trials (Trials 6–8, 11, 16, 17 in Table 1) were of interventions initiated in the second trimester to prevent preterm birth in high risk women. For all of these trials, the MFMU Network has chosen an endpoint based on preterm birth before a specific gestational age cutoff. In contrast, some study groups have used neonatal morbidity or mortality as the primary outcome, almost always as a composite. Their argument is that preterm birth is a surrogate for the morbidity and mortality experienced by a premature infant.⁷ But it could also be argued that neonatal outcome is a surrogate for adverse health or disability after neonatal discharge. Therefore, using a preterm birth cutoff as the primary outcome is logical since the very structure of the intervention, such as a pessary, is to prevent preterm birth from happening. If a trial is positive, however, it is

possible that fetuses will be exposed to a new intervention as standard of care, so that conducting a long term follow-up of the children would be important.

It is important, however to pick an appropriate gestational age cutoff for the population being studied – we have chosen 35 weeks for twin gestation and 37 weeks for singleton gestation. It should be noted that in all of these trials, fetal demise occurring before the cutoff is also included in the outcome. We found that this needed to be specified especially clearly in the case of multifetal gestation, as it is possible for example, to have a fetal loss in one twin while the other twin survives and may be born days or even weeks later. Of note, preterm birth not only satisfies all of the criteria for a good primary outcome, but it is probably the easiest of all outcomes to obtain. All that is needed is a standardized estimate of gestational age from a pre- randomization ultrasound and the delivery date.

Three of the trials (1, 2, 12 in Table 1) are of interventions intended to ameliorate the effects of premature birth in the baby when preterm delivery appears to be imminent. In the BEARS trial⁸ (2 in Table 1) of repeated versus single dose corticosteroids, the primary outcome was a composite of neonatal morbidities including respiratory outcomes that are common in very premature babies and had been shown to be responsive to a single course of antenatal steroids. Of note, there was considerable discussion as to whether a safety secondary outcome should be elevated to the status of ‘primary’ since there were growing concerns regarding the effect of repeated steroids on fetal growth.⁹ Ultimately it was decided that a single efficacy endpoint was most appropriate, as that was the main question to be answered. However, we included neonatal anthropometric measures obtained with standardized equipment by trained research as major secondary outcomes in the protocol. We also conducted a follow-up study of the children at age 2 to 3 years, which included anthropometry.¹⁰

In the ALPS trial (trial 12 in Table 1) of corticosteroids for women at high risk of delivering in the late preterm period, the primary outcome was a composite endpoint describing the need for respiratory support in the first three days of life.¹¹ Follow-up of these children at 6 years of age is planned.

In each of the three trials in this group, the relatively complex primary outcomes required neonatal, neurodevelopmental or pediatric expertise beyond that of the MFMU trialists for outcome determination which usually takes the form of a blinded review of the neonatal chart or standardized examination of the infant. The need for a multidisciplinary approach is a common characteristic of MFMU Network trials and other pregnancy trials where the main focus is the baby rather than the mother.

Three trials (trials 9, 10, 14 in Table 1) involved screening for and treating a maternal medical condition that could potentially have adverse consequences for the infant, but otherwise the mother would not normally need treatment. For one of these, the CMV trial (14), there was a discussion regarding the appropriate choice of primary outcome. The purpose of this ongoing trial is to evaluate monthly infusions of hyper-immune globulin versus placebo as an intervention for pregnant women who have been exposed to the cytomegalovirus (CMV) during pregnancy. Fetuses exposed to CMV in utero that acquire

the infection are at high risk for death, hearing loss, chorioretinitis, neurodevelopmental delay and other adverse outcomes as children, especially if symptomatic in-utero or at birth. The choice of primary outcome was not completely straightforward. Some argued that the primary outcome should be based on adverse outcome at 2 years; however, there was concern that this would result in an infeasible sample size. The Network chose congenital infection as the primary outcome. Nevertheless, children are being followed up and a secondary outcome was constructed consisting of ordered categories (not infected, infected but without severe disability, infected with severe disability, death). Assuming an underlying trend in the hyper-immune globulin group of a reduction in severity of outcome across categories, there is sufficient power to show a difference with the planned sample size which would not have been the case if we had used the dichotomous outcome of disability or death..

The Network has three trials of obstetric management at term (trials 3, 13, 15 in Table 1). For each of these studies, fetal benefit is essentially balanced against avoiding excess cesarean delivery. Two of the trials involved a device to be used as an adjunct to electronic fetal monitoring so that the care provider has more information on which to base a delivery decision. In the first trial (FOX, trial 3 in the table), a fetal pulse oximeter was tested.¹² If the fetal oxygen saturation was adequate even in the face of a non-reassuring (but not ominous) fetal heart rate pattern, the managing physician was supposed to continue labor. However, if the fetal oxygen saturation was inadequate the physician was expected to manage the labor based on the fetal heart rate patterns as they normally would without knowing fetal oxygen saturation. Thus the purpose was to improve the specificity but not the sensitivity of electronic fetal monitoring – i.e. to reduce the number of unnecessary cesarean deliveries. It followed that cesarean delivery would be the primary outcome. The safety concern was that delaying delivery might increase adverse neonatal outcome. A logical option might have been to design the trial to test whether neonatal outcome was the same or only marginally worse (non-inferior) in the fetal pulse oximetry group compared with those who had electronic fetal monitoring alone as a joint primary question. We did not, and it is a moot question as to whether we should have, as there was no difference in cesarean delivery rates and the technology is no longer being used.

In the second trial (STAN, trial 13 in the table) a proprietary monitor (STAN monitor) was tested which, in addition to the normal function of an electronic fetal heart rate monitor, analyzes fetal ECG information reflective of myocardial metabolism and acid-base balance.¹³ The STAN monitor issues a visual alert – known as a STAN event – when changes occur. If a STAN event occurs with an ‘uncertain’ fetal heart rate pattern (essentially neither reassuring nor ominous), the physician is supposed to attempt intrauterine resuscitation to improve the fetal condition or deliver expeditiously. On the other hand, if no STAN event occurs in the face of an uncertain fetal heart rate pattern, the provider could be reassured and allow labor to continue. Unlike the fetal pulse oximeter, this device was intended to improve both sensitivity and specificity of electronic fetal heart rate monitoring. For this trial therefore a neonatal composite endpoint was chosen representing a cluster of neonatal outcomes that manifest early, indicate that the fetus may have been compromised during labor, that are associated with risk of long term neurological adverse

outcomes, and potentially could be avoided by more prompt delivery.¹⁴ Cesarean delivery was a secondary outcome.

The third trial in this group is ARRIVE (trial 15 in table 1), a randomized trial of elective induction at 39 weeks versus expectant management in nulliparous women. The primary outcome is a neonatal composite, somewhat similar to that used for the STAN trial above, with additional components representing adverse events that may be increased in post term pregnancies. One of the fears about elective induction at 39 weeks is that it may lead to an increase in cesarean deliveries. In this case, we did consider cesarean delivery as a non-inferiority co-primary endpoint so that if there was neonatal benefit in the induction group while the cesarean rate was at most only a bit worse, elective induction would be considered superior. However, we concluded that women and their providers would be willing to accept an increased risk of cesarean delivery to avoid the adverse neonatal outcome. Moreover, the concern regarding the potential increase in the cesarean rate was based on studies comparing electively-induced women with those who labored spontaneously rather than with those expectantly managed, some of whom will also require induction or even a cesarean without labor. More recent studies using the correct comparison group appear to show a decrease in the cesarean rate with elective induction, if anything.

Composite Outcomes

Composite outcomes are constructed as a combination of individual outcomes to create a single endpoint. Most often it is a binary endpoint such that if any of the individual component outcomes have occurred, the composite outcome is deemed to have occurred. From the previous discussion and Table 1, it can be seen that all except one of the primary endpoints in Network trials considered here are composite outcomes of some form or other. An often cited disadvantage of a composite is the potential difficulty of interpretation especially when outcomes of different severity are combined or the treatment effect on the different components are in different directions.¹⁵ However, the use of composite neonatal outcomes is particularly widespread and a natural choice in pregnancy trials as neonates, especially those born prematurely, are subject to a constellation of potential adverse outcomes. An example is the BEARS trial where the primary outcome was a composite of perinatal death, severe respiratory distress syndrome (RDS), chronic lung disease, intraventricular hemorrhage (IVH) grade III or IV, and periventricular leukomalacia (PVL).

Advantages of composite outcomes include a smaller sample size, avoiding an arbitrary choice between multiple relevant outcomes and mitigation of competing risks. A competing risk is an event that precludes observation of the event of interest. Competing risks are particularly prevalent in pregnancy trials since fetal death is a competing risk for any neonatal outcome. By including the competing risk as a component along with the outcome(s) of interest in a composite, the resulting outcome is then measurable in all participants, one of the criteria for an appropriate primary outcome. In the BEAM trial of antenatal magnesium sulfate to prevent cerebral palsy (CP), we included fetal or infant death prior to one year of age rather than two years in the primary outcome as a competing risk to moderate or severe cerebral palsy at age 2. This was because infants were also evaluated at

one year of age for CP – thus if they died between one and two years of age, we had already had the opportunity to determine whether they had CP.

Spontaneous preterm birth is another example of an outcome that has a competing risk. However, indicated preterm delivery is a competing risk because once that occurs, a woman cannot have a spontaneous preterm birth. For this reason, we always include all preterm births, whether iatrogenic or spontaneous, in our primary outcomes.

It is important to take the competing risk component into consideration in the sample size estimate. If it is not expected that the intervention will affect the rate of the competing outcome, the overall effect size to be detected should be decreased accordingly. It is also important that the assumed component event rates for the sample size calculation be close to accurate. In the BEAM trial, we assumed that there would be an approximately 50% reduction in CP, but no reduction in death. The observed effect sizes were fairly close to those assumptions. However, our estimate of the expected rate of moderate to severe cerebral palsy in the sample size calculation was much higher than observed (more than double) and our estimate of death before one year of age was about 30% lower than observed. As a result, there was no significant difference in the primary outcome between the two treatment groups.

Bias

Bias is a systematic error that distorts the truth. In clinical research, there are many sources of bias, from before enrollment, during study conduct, during analysis and even beyond. Selection bias is the differential selection of participants and/or their outcomes. Randomization properly carried out protects against differential selection of the participants. A valid primary outcome that adheres to the criterion ‘measurable in all participants’ avoids one potential source of selection bias. However, loss to follow up also potentially creates selection bias. In the MFMU Network, we emphasize the importance of obtaining as good a follow-up rate as possible; we set aggressive goals and monitor performance. Salazar et al. in this issue describe some of the methods used to minimize loss to follow-up particularly for the trials where the primary outcome is determined after neonatal discharge.

Information bias is a distortion in the relationship between intervention and outcome. The best protection against information bias in a randomized trial is the double-blind design. Most of the MFMU trials have been double blind, but we have conducted or are conducting three completely unblinded trials (FOX, STAN, ARRIVE) and two with partial blinding (GDM, PROSPECT). When double blinding is not possible, there is a risk that the outcome will be assessed differently according to which group the participant has been assigned. Bias in this case may be minimized by choosing a ‘hard’ endpoint that is not determined by clinical practice or by the behavior of the participant – for example, neonatal death as opposed to interventions used in the neonatal intensive care unit, or a diagnosis rather than maternal report of symptoms. It is also minimized by having independent reviewers blinded to treatment assignment assess the primary outcome and other outcomes. In the STAN trial, the primary outcome was defined by objective criteria and the outcomes were reviewed by a

small group of MFMU Network investigators blinded to randomization assignment. Of note, no investigator was allowed to review outcome from his/her own center.

Also in the STAN trial, women were assigned either to 'open' or 'masked' fetal ST analysis monitoring at randomization. If assigned to 'open' monitoring, the STAN machine functioned as intended with STAN events displayed when they occurred. In the masked monitoring mode, the STAN monitor software automatically suppressed the STAN events, and the machine functioned as a conventional fetal heart rate monitor. Although women were not necessarily completely blinded to the type of monitoring, they were treated very similarly between the two groups. This helped to reduce information bias. The same idea was used in the FOX trial

Contamination is another form of information bias. In this scenario, those assigned to the intervention are treated or behave more like those assigned to control and/or vice versa. For example, in a trial of a new type of physician management versus a standard of care, physicians may ignore the intervention and manage patients by standard of care like the control group. This could lead to a self-fulfilling prophecy of no benefit although the new treatment may actually be effective. This was a concern in STAN trial. To overcome the potential for contamination, we instituted a rigorous policy of training and certification for obstetric providers and a group of investigators masked to outcome monitored obstetric management to ensure that they were following manufacturer's guidelines in the open STAN arm.

Contamination was also a concern when designing the GDM trial (trial 4 in Table 1). The purpose of the trial was to determine whether women with mild GDM (fasting glucose < 95 mg/dl) should be treated with usual GDM care, i.e. diet, glucose monitoring, and insulin as needed, as opposed to normal obstetric care.¹⁶ Usual GDM care had evolved as a standard of care without good evidence. This trial was intended to challenge the current dogma. A previous pilot trial had attempted to do the same thing, but the null results were called into question because it was felt that in the control arm women could have been 'self-treating' by modifying their own diet on the basis of self-education.¹⁷ To avoid this possibility, as well as the potential for caregivers to modify practice for women with mild GDM randomized to normal obstetric care, we used a partial blinded design. In the Network trial, OGTT tests were carried out at a central laboratory and results forwarded to the data coordinating center. Women who were designated as having mild GDM according to their OGTT results were randomized by the coordinating center to usual GDM care or to standard obstetric care. A matching sample of women who passed the OGTT test was also enrolled by the coordinating center into the usual obstetric care arm. The clinical center staff were notified that either a woman was randomized to usual GDM care or had been enrolled in the usual obstetric care group. However, the staff were not aware of which of those in the obstetric care group had mild GDM and were part of the randomized trial and which had a negative OGTT test and were enrolled mainly to mask the identity of women in the trial. Likewise, the women who were assigned to usual obstetric care knew that they were in the study, but did not know whether they had mild GDM or not until after their participation ended.

Discussion

There are many challenges in conducting trials in pregnancy and this report highlights just a few. In the previous issue of this journal in which we described some of the lessons learned from conducting trials in the Network, we focused mainly on the reasons for recruitment difficulties including lack of equipoise and/or buy-in from the investigators or referring physicians and difficulty in obtaining good estimates of outcome rates for sample size calculations. While these difficulties have not retreated, we are perhaps better equipped to anticipate and deal with them. Therefore we have focused here on design issues – the first and foremost of which is defining the primary question to be answered. This is greatly affected by the perceived balance of risk and benefit between mother and baby, and in turn informs the choice of the primary and secondary outcomes.

Defining the outcomes to be interpretable is critical, and as a group we spend considerable time in discussion. The choice of a good primary outcome is an area where we have learned lessons from one trial to apply to a future trial, since we have conducted several groups of trials with similar goals (preventing preterm birth, ameliorating the effects of preterm birth, screening and treating women for a condition that could have an important effect on the baby, and determining when to deliver). Composite neonatal outcomes are used often in our trials and other trials in pregnancy. However, it is important that different trial groups define outcomes consistently and some progress on that is being made.¹⁸

There are many areas where bias can undermine the interpretability of the results. A good choice of primary outcome and excellent follow-up are important in all trials regardless of blinding. In unblinded trials, there is a particular risk of bias. We have learned that it is important to address this upfront and to keep the investigators blinded as far as possible. We use hard endpoints wherever possible, central review of outcomes, and monitor protocol adherence. We try to conduct the trials rigorously, but without making them overcomplicated or collecting unnecessary data. Overall our goal is to get valid answers to important clinical questions that affect pregnant women and their babies.

References

1. Lyerly AD, Little MO, Faden R. The second wave: Toward responsible inclusion of pregnant women in research. *Int J Fem Approaches Bioeth.* 2008; 1:5–22. [PubMed: 19774226]
2. Grimes, D. Discovering the need for randomized controlled trials in obstetrics: a personal odyssey. (Accessed at www.jameslindlibrary.org.)
3. Goldkind SF, Sahin L, Gallauresi B. Enrolling pregnant women in research--lessons from the H1N1 influenza pandemic. *N Engl J Med.* 2010; 362:2241–2243. [PubMed: 20554981]
4. Grant A. Rationale for and work of the Perinatal Trials Service. *Early Hum Dev.* 1992; 29:305–308. [PubMed: 1396257]
5. Wright LL, McNellis D. National Institute of Child Health and Human Development (NICHD)-sponsored Perinatal Research Networks. *Semin Perinatol.* 1995; 19:112–123. [PubMed: 7604302]
6. Thom EA, Rouse DJ. National Institute of Child H, Human Development Maternal-Fetal Medicine Units N. What we have learned about conducting randomized controlled trials in the NICHD MFMU network. *Semin Perinatol.* 2003; 27:253–260. [PubMed: 12889593]
7. Meher S, Alfirevic Z. Choice of primary outcomes in randomised trials and systematic reviews evaluating interventions for preterm birth prevention: a systematic review. *BJOG.* 2014; 121:1188–1194. discussion 95–6. [PubMed: 24571433]

8. Wapner RJ, Sorokin Y, Thom EA, et al. Single versus weekly courses of antenatal corticosteroids: evaluation of safety and efficacy. *Am J Obstet Gynecol.* 2006; 195:633–642. [PubMed: 16846587]
9. Bonanno C, Fuchs K, Wapner RJ. Single versus repeat courses of antenatal steroids to improve neonatal outcomes: risks and benefits. *Obstet Gynecol Surv.* 2007; 62:261–271. [PubMed: 17371606]
10. Wapner RJ, Sorokin Y, Mele L, et al. Long-term outcomes after repeat doses of antenatal corticosteroids. *N Engl J Med.* 2007; 357:1190–1198. [PubMed: 17881751]
11. Gyamfi-Bannerman, C.; Thom, EA.; Blackwell, SC., et al. Antenatal Betamethasone for Women at Risk for Late Preterm Delivery. (Accessed at <http://www.nejm.org/doi/full/10.1056/NEJMoa1516783>.)
12. Bloom SL, Spong CY, Thom E, et al. Fetal pulse oximetry and cesarean delivery. *N Engl J Med.* 2006; 355:2195–2202. [PubMed: 17124017]
13. Belfort MA, Saade GR, Thom E, et al. A Randomized Trial of Intrapartum Fetal ECG ST-Segment Analysis. *N Engl J Med.* 2015; 373:632–641. [PubMed: 26267623]
14. Executive summary: Neonatal encephalopathy and neurologic outcome, second edition. Report of the American College of Obstetricians and Gynecologists' Task Force on Neonatal Encephalopathy. *Obstet Gynecol.* 2014; 123:896–901. [PubMed: 24785633]
15. Ross S. Composite outcomes in randomized clinical trials: arguments for and against. *Am J Obstet Gynecol.* 2007; 196:119 e1–119 e6. [PubMed: 17306647]
16. Landon MB, Spong CY, Thom E, et al. A multicenter, randomized trial of treatment for mild gestational diabetes. *N Engl J Med.* 2009; 361:1339–1348. [PubMed: 19797280]
17. Garner P, Okun N, Keely E, et al. A randomized controlled trial of strict glycemic control and tertiary level obstetric care versus routine obstetric care in the management of gestational diabetes: a pilot study. *Am J Obstet Gynecol.* 1997; 177:190–195. [PubMed: 9240606]
18. van 't Hooft J, Duffy JM, Daly M, et al. A Core Outcome Set for Evaluation of Interventions to Prevent Preterm Birth. *Obstet Gynecol.* 2016; 127:49–58. [PubMed: 26646133]
19. Rouse DJ, Hirtz DG, Thom E, et al. A randomized, controlled trial of magnesium sulfate for the prevention of cerebral palsy. *N Engl J Med.* 2008; 359:895–905. [PubMed: 18753646]
20. Roberts JM, Myatt L, Spong CY, et al. Vitamins C and E to prevent complications of pregnancy-associated hypertension. *N Engl J Med.* 2010; 362:1282–1291. [PubMed: 20375405]
21. Rouse DJ, Caritis SN, Peaceman AM, et al. A trial of 17 alpha-hydroxyprogesterone caproate to prevent prematurity in twins. *N Engl J Med.* 2007; 357:454–461. [PubMed: 17671253]
22. Caritis SN, Rouse DJ, Peaceman AM, et al. Prevention of preterm birth in triplets using 17 alpha-hydroxyprogesterone caproate: a randomized controlled trial. *Obstet Gynecol.* 2009; 113:285–292. [PubMed: 19155896]
23. Harper M, Thom E, Klebanoff MA, et al. Omega-3 fatty acid supplementation to prevent recurrent preterm birth: a randomized controlled trial. *Obstet Gynecol.* 2010; 115:234–242. [PubMed: 20093894]
24. Casey B. Effect of treatment of maternal subclinical hypothyroidism or hypothyroxinemia on IQ in offspring. *Am J Obstet Gynecol.* 2016; 214:S2-S.
25. Grobman WA, Thom EA, Spong CY, et al. 17 alpha-hydroxyprogesterone caproate to prevent prematurity in nulliparas with cervical length less than 30 mm. *Am J Obstet Gynecol.* 2012; 207:390 e1–390 e8. [PubMed: 23010094]

Ongoing and Completed Randomized Trials Conducted by the MEMU Network since 2003

Table 1

Short Name	Intervention(s)	Control	Population (risk)	Sample Size	Primary Outcome
1 BEAM ¹⁹	Magnesium sulfate infusion	Placebo infusion	At high risk of preterm delivery at 24–31 weeks	2,241	Fetal/infant death < 1 year or moderate to severe CP at age 2
2 BEARS ⁸	Weekly courses of betamethasone	Weekly courses of placebo	At high risk for preterm delivery at 23–31 weeks, Single course of cortico-steroids 1 week earlier	495	Composite of perinatal death; severe RDS; chronic lung disease; IVH grade III or IV; PVL
3 FOX ¹²	Fetal pulse oximeter + conventional EFM	EFM alone	Nulliparous, singleton in early labor, 36 weeks	5,341	Cesarean delivery
4 GDM ¹⁶	Usual gestational diabetes (GDM) care (diet, glucose monitoring, insulin if needed)	Usual obstetric care	Singleton, mild GDM, 24–30 wks	958	Composite of perinatal death; neonatal hypoglycemia; hyperinsulinemia; hyperbilirubinemia; birth trauma
5 CAPPS RCT ²⁰	Antioxidants (vitamins E and C)	Placebo	Nulliparous, singleton normotensive, 9–16 wks	10,154	Composite of perinatal death; severe hypertension with thrombocytopenia, elevated serum creatinine, or elevated liver enzymes; eclampsia; indicated delivery for hypertension < 32 weeks;
6 STTARS (twins) ²¹	17 OHP	Placebo	Twin gestation 16–20 wks	661	SGA < 3 rd percentile Preterm delivery or fetal death < 35 weeks
7 STTARS (triplets) ²²	17 OHP	Placebo	Triplet gestation, 16–20 wks	134	Preterm delivery or fetal death < 35 weeks
8 Omega-3 ²³	Omega-3 fatty acid supplement + open labeled 17OHP	Placebo + open labeled 17OHP	Previous spontaneous preterm delivery, singleton, 16–21 wks	852	Preterm delivery or fetal death < 37 weeks
9 TSH-S ²⁴	Thyroxine	Placebo	Subclinical hypothyroid (high TSH with normal free T4), 8–20 weeks	677	IQ at age 5 or DAS at age 3 or death < 3 years
10 TSH-H ²⁴	Thyroxine	Placebo	Hypothyroxinemia (low free T4 with normal TSH), 8–20 weeks	526	IQ at age 5 or DAS at age 3 or death < 3 years
11 SCAN ²⁵	17 OHP	Placebo	Nulliparous, singleton, short cervix < 30 mm, 16–22 weeks	657	Preterm delivery or fetal death < 37 weeks

Short Name	Intervention(s)	Control	Population (risk)	Sample Size	Primary Outcome
12 ALPS ¹¹	Course of betamethasone	Placebo	Singleton, at high risk of preterm delivery 34–36 weeks	2,831	Composite of CPAP or HFNC for at least 2 hours, 30% oxygen for at least 4 hours, ventilator use, ECMO, stillbirth, or death in first 72 hours
13 STAN ¹³	Fetal ECG analysis of the ST segment + EFM	EFM alone	Singleton, in labor, 36 weeks	11,108	Composite of perinatal death, HIE, acidosis, intubation, 5-min Apgar score <4, seizure
14 CMV (ongoing)	CMV hyperimmune globulin	Placebo	Singleton, CMV positive, <24 weeks	800	CMV positive fetus or neonate; fetal death or neonatal death
15 ARRIVE (ongoing)	Elective induction at 39 weeks	Expectant management	Singleton, nulliparous, 38 weeks	6,000	Composite of perinatal death, HIE, CPAP, intubation or HFNC by 72 hrs, 5-min Apgar score <4, seizure, MAS, sepsis, pneumonia, birth trauma, intracranial hemorrhage, hypotension
16 PROSPE CT (ongoing)	1) Arabin pessary 2) Vaginal progesterone	Placebo for vaginal progesterone	Twin gestation, short cervix <30 mm, 16–23 weeks	630	Preterm delivery or fetal death <35 weeks
17 APPS* (to start)	Arabin pessary + open labeled vaginal progesterone	Vaginal progesterone	Singleton gestation, short cervix 20 mm, 16–23 weeks	850	Preterm delivery or fetal death <37 weeks

*. Treatment groups may change

17-OHP – 17 α -hydroxyprogesterone caproate
 HFNC – high flow nasal cannula
 MAS – meconium aspiration syndrome
 EFM – electronic fetal monitoring
 CPAP – continuous positive airway pressure
 HIE – hypoxic ischemic encephalopathy