17α-Alpha Hydroxyprogesterone Caproate for Prevention of Preterm Birth

Overview of FDA Background Document

Introduction

Adeza Biomedical has submitted New Drug Application (NDA) 21-945 for 17α-hydroxyprogesterone caproate (17OHP-C) injection for the proposed indication:

> "Prevention of preterm birth in pregnant women with a history of at least one spontaneous preterm birth"

Preterm birth is defined as a birth prior to 37 weeks gestational age.

The proposed dosing regimen is a weekly intramuscular injection of 250 mg of 17OHP-C in 1 mL castor oil with 46% benzyl benzoate and 2% benzyl alcohol, beginning at 16 weeks 0 days (16^0) to 20 weeks 6 days (20^6) weeks gestation and used through 36^6 weeks gestation or birth.

Currently there is no drug product approved in the United States for prevention of preterm birth; however, 17OHP-C is being compounded by pharmacists and is being used widely for prevention of preterm birth in women at high risk. The medical need for an approved drug product for prevention of preterm birth is particularly acute because there also are no approved drug products currently marketed in the United States for the treatment of preterm labor. Although several drug products with tocolytic properties (i.e., stopping uterine contractions) are used off-label for treatment of preterm labor, randomized controlled trials have failed to demonstrate that these drugs improve perinatal outcomes.

In 2003, the findings from a multicenter, randomized, placebo-controlled, double-blind clinical trial of 17OHP-C in women at high risk for preterm birth were published. This trial was sponsored by the National Institute for Child Health and Human Development (NICHD) and was conducted by the Maternal-Fetal Medicine Units (MFMU) Network, which at that time consisted of approximately 19 university-based clinical centers in the U.S. This study (referred to as Study 17P-CT-002 in this document) showed a 34% reduction in preterm births prior to 37^0 weeks in women with a prior preterm birth (a population at high risk for a recurrent preterm birth).

NDA 21-945 is based largely on the clinical data from Study 17P-CT-002 and a follow-up study to support the safety and effectiveness of 17OHP-C for the prevention of preterm birth. The database submitted by the Applicant to support safety and effectiveness includes data from the following three studies:

- Initial Formulation Study (Study 17P-IF-001**).** This study began in February 1998, and 150 of the proposed 500 subjects were randomized. Treatment was terminated in March 1999 because the active study drug (17OHP-C) was recalled by its manufacturer, under the direction of the FDA, due to violations of manufacturing practices. Eighty-six subjects completed the treatment regimen before the study was stopped: 57 (61%) of the 17OHP-C subjects and 29 (52%) of the placebo subjects.
- Primary Clinical Trial for Safety and Efficacy (Study 17P-CT-002). This study, which was started in October 1999, randomized 463 subjects who had at least one documented prior spontaneous preterm birth of a singleton, non-anomalous fetus. Of these, 418 subjects (90.3%) completed dosing through 36^6 weeks or birth: 279 (90.0%) in the

17OHP-C group and 139 (90.8%) in the placebo group. This study was terminated prior to enrolling the planned 500 subjects because the pre-specified stopping criterion for efficacy was attained at an interim analysis.

- Follow-up Study of the Children from the 17P-CT-002 Trial (Study 17P-FU). This was a follow-up to Study 17P-CT-002. The follow-up study collected data with a validated child development instrument, the Ages and Stages Questionnaire (ASQ), a Survey Questionnaire concerning the health and development of the child, and a physical examination. The children were at least 2 years of age at the time of the follow-up assessments. The primary objective of this study was to determine whether there was a difference in achievement of developmental milestones and physical health between children born to women who received weekly intramuscular injections of 17OHP-C compared with placebo during the pregnancy in Study 17P-CT-002.

Points for the Advisory Committee to Consider

The major issues that the FDA would like the Advisory Committee for Reproductive Health Drugs to consider include:

Adequacy of Clinical Data to Support the Effectiveness of 17OHP-C

In general, the FDA requires an Applicant for a new drug product to submit two adequate and well-controlled clinical trials as substantial evidence of effectiveness. One of the circumstances in which a single clinical trial may be used as substantial evidence of effectiveness is a trial that has demonstrated a clinically meaningful effect on mortality, irreversible morbidity, or prevention of a disease with potentially serious outcome, and confirmation of the result in a second trial would be logistically impossible or ethically unacceptable.

The Applicant is seeking approval for 17OHP-C based primarily on (1) the findings from a single clinical trial and (2) a surrogate endpoint for neonatal/infant morbidity and mortality (i.e., reduction in the incidence of preterm births at less than 37 weeks gestation).

Although preterm birth is defined as a birth prior to 37 weeks gestation, the clinical significance of preterm birth is more pronounced prior to 32 weeks gestation. In the U.S., infants born after 32 weeks have very low mortality rates, and relatively low long-term morbidity. However, since a larger number of preterm births occur after 32 weeks gestation, the public health importance of preventing even these later gestational age preterm births may be noteworthy.

Study 17P-CT-002 demonstrated a statistically significant reduction in the primary endpoint of preterm births prior to 37^0 weeks gestation. However, the reduction in preterm births prior to 35^0 weeks and prior to 32^0 weeks gestation, better surrogates for significant neonatal morbidity or mortality, was not statistically persuasive. In addition, the primary clinical trial did not demonstrate a significant reduction in another secondary endpoint, a composite assessment of infant mortality and morbidity.

The FDA asks the Advisory Committee whether the primary endpoint, prevention of preterm birth prior to 37 weeks, is an adequate surrogate for infant mortality and morbidity. If so, does the available information provide sufficient evidence of effectiveness such that an additional confirmatory clinical trial is not warranted?

Generalizability of Efficacy Results

The results of Study 17P-CT-002 demonstrate a reduction in the rate of preterm birth prior to 37 weeks from the **55%** incidence seen in the placebo group to the 36% incidence observed in the 17OHP-C group. However, a previous large clinical trial sponsored by the NICHD (on which the sample size calculations for the current clinical trial were based) found the incidence of preterm birth prior to 37 weeks in an untreated, but similarly high risk population to be **37%**. The incidence of preterm births in the placebo arm of Study 17P-IF-001 (also conducted by the MFMU Network) was **36%**.

The FDA asks the Advisory Committee whether the difference in the incidence of preterm birth prior to 37 weeks observed in the placebo group of this trial as compared to another MFMU Network trial evaluating a similar untreated high risk population suggests the need to replicate the findings of Study 17P-CT-002 in a confirmatory study. Does the Committee believe that the efficacy findings of Study 17P-CT-002 would be applicable to women in the general U.S. population who have a history of one or more preterm births?

Potential Safety Signal

There was a trend toward an increase in second trimester miscarriage rate (pregnancy loss prior to 20 weeks' gestation) and a suggestion of an increase in stillbirth rate (death of the fetus prior to or during delivery) in the 17OHP-C group.

The FDA asks the Advisory Committee whether further studies are needed to evaluate the potential association of 17OHP-C with increased risk of second trimester miscarriage and stillbirth.

Gestiva

(17α-hydroxyprogesterone caproate)

NDA 21-945

Proposed Indication

"GESTIVA is indicated for the prevention of preterm birth in pregnant women with a history of at least one spontaneous preterm birth"

Dosing Regimen

GESTIVA is to be administered intramuscularly at a dose of 250 mg (1 mL) once each week beginning at 16 weeks 0 days (16^0 weeks) to 20 weeks 6 days (20^6 weeks) of gestation to week 37 of gestation or until birth.

Drug Product

GESTIVA will be supplied as 5 mL of a sterile solution in a multiple dose glass vial. Each mL will contain 17α-hydroxyprogesterone caproate USP, 250 mg/mL (25% w/v), castor oil (28.6% v/v), benzyl benzoate (46% v/v), and benzyl alcohol (2% v/v) as preservative.

Review by the Division of Reproductive and Urologic Products
Food and Drug Administration
August 2, 2006

Table of Contents

Table of Tables

Table of Figures

APPENDICES

APPENDIX 1 GUIDANCE FOR INDUSTRY: PROVIDING CLINICAL EVIDENCE OF EFFECTIVENESS FOR HUMAN DRUG AND BIOLOGICAL PRODUCTS, MAY 1998

APPENDIX 2 SURVEY QUESTIONNAIRE

APPENDIX 3 AGES AND STAGES QUESTIONNAIRE A PARENT-COMPLETED, CHILD-MONITORING SYSTEM

APPENDIX 4 PUBLICATIONS

1 BACKGROUND

1.1 Public Health Significance of Prematurity

Preterm birth (PTB), birth prior to 37 weeks of gestational age, is the leading cause of neonatal mortality (infant death <28 days of life) and is a major cause of early childhood mortality and morbidity in the United States.[1] As many as half of all pediatric neurodevelopmental problems can be attributed to preterm birth.[2] The U.S preterm birth rate increased by 29% over the previous 2 decades to a high of 12.1% in 2002.[3] Most of this increase occurred in preterm births of 32-36 weeks gestational age and is thought to be due to the increasing frequency of pregnancy in women older than 35 years and the use of infertility treatments.[4] The rate for very early preterm births (< 32 completed weeks gestation) has remained stable at about 2% of all births; however, most perinatal/neonatal and infant mortality/morbidity occurs in these infants.[3] Preterm births most often result from spontaneous preterm labor and preterm premature rupture of membranes (pPROM).[5,6,7] However, 20-30% of preterm births are considered "indicated" to avoid or minimize maternal/fetal complications.[8]

Rates of PTB in the United States differ profoundly among ethnic groups; the rate of PTB in non-Hispanic black births is twice as high as that of non-Hispanic white births. These disparities remain even after adjusting for confounders such as education and occupation, suggesting a combination of genetic, environmental, and social factors as the etiology.[9,10,11,12,13,14]

Accurate prediction and prevention of PTB remain elusive.[2,6-8,15-19] Most biomarkers to assess the risk of PTB have poor positive predictive value to guide clinical decisions.[2,8,15-20] Examples of risk factors include history of previous preterm birth; multifetal gestation; and cervical, uterine, and placental structural or physiologic abnormalities.

Prophylactic methods for prevention of preterm birth, including drugs, bed rest, or other interventions, have been shown in general to lack effectiveness. Tocolytic drugs may be given to reduce the frequency of uterine contractions. However, they have not been efficacious in preventing preterm birth nor have they resulted in improved newborn outcomes.

Preterm birth has been described as a "common, complex disorder, stemming from heterogeneous composites of multiple gene-environment interactions."[21] Evidence supporting this includes findings of familial aggregation, non-Mendelian heritability, high rates of recurrence, and the existence of ethnic/racial disparities.

1.2 Description and Causes of Prematurity

The "syndrome" of PTB is now understood as the clinical endpoint for a number of potential causes. Four major pathophysiologic pathways have been hypothesized:

(1) inflammation/infection with its associated maternal and fetal cytokine response

(2) maternal/fetal stress with generation of placental and fetal membrane-derived corticotropin-releasing hormone, which enhances placental estrogen and fetal adrenal cortisol production

(3) abruption or decidual hemorrhage with thrombin-induced protease expression and disturbances in uterine tone

(4) mechanical stretch due to multifetal pregnancy or polyhydramnios-induced abnormal uterine and cervical distension

Infection/inflammation is the only pathologic process for which a firm causal link with prematurity has been established and for which a defined molecular pathophysiology is known.[22] It has been estimated that 40% of all preterm births occur to mothers with intrauterine infection, which is usually subclinical. The lower the gestational age at delivery, the greater the frequency of intrauterine infection.[23] The most common pathway is ascending organisms from the lower genital tract, more commonly from an alteration in the normal vaginal flora.[24] The organisms enter the amniotic cavity and then, in some cases, will gain access to the fetus which may result in fetal sepsis or the Fetal Inflammatory Response Syndrome (FIRS).[25] The clinician managing preterm labor must balance the possibility of sub-clinical infection, against the sequelae of prematurity, both having the potential for causing death.

1.3 Clinical History and Background Data on 17α-hydroxyprogesterone Caproate

17α-hydroxyprogesterone caproate (17OHP-C) was approved by the Food and Drug Administration (FDA) in 1956 for use in pregnant women (NDA 10-347; Delalutin®). The approved indications included the treatment of habitual and recurrent abortion, threatened abortion, and post-partum "after pains." This approval was based largely on safety consideration in that it occurred prior to the FDA Drug Amendment of 1962, which required that drugs must have substantial evidence of efficacy in addition to evidence of safety in adequate and well-controlled trials. In 2000, the FDA withdrew approval for Delalutin. This action was taken at the request of the holder of the NDA because the holder was no longer marketing the drug. The action was not taken because of safety concerns.

The published literature includes several studies evaluating the efficacy of 17OHP-C in preventing preterm birth (see Table 1). Not included in Table 1 is the publication by Meis PJ, Klebanoff M, et al. that was based on the finding from Study 17P-CT-002 (the primary study supporting the efficacy and safety of 17OHP-C in this NDA.)

Table 1 Studies of the Efficacy and Safety of 17OHP-C in Preventing Preterm Birth

Investigator	Drug:Dose	Entry Criteria	Design	Subjects	Start	Stop	Outcome % PTB[A]	No. of SAB[B]
LeVine 1964[1]	17P: 500 mg weekly vs. Placebo	3 SABs[B]	RCT, DB[C] Placebo 1:1	17P: 15 Placebo: 15	< 16 wks	36 wks	17P: 7/15 (46%) Placebo: 10/15 (66%)	17P: 3/15 Placebo: 7/15
Papiernik-Berkhauer 1970[2]	17P: 250 mg q 3 days vs. Placebo	High preterm risk score	RCT Placebo 1:1	17P: 50 Placebo: 49	28 – 30 wks	8 doses	17P: (4.1%) Placebo: (18.8%)	
Johnson et al 1975[3]	17P: 250 mg weekly vs. Placebo	2 SABs[B] or 1PTB[A] + 1 SAB[B] or hx 2 PTBs	RCT, DB[C] Placebo 1:1	17P: 18 (4 cerclage) Placebo: 22 (3 cerclage)	Booking < 24 wks	37 wks	17P: 0/18 (0%) Placebo: 9/22 (41%)	17P: 3/23 Placebo: 0/27
Yemini 1985[4]	17P: 250 mg weekly + cerclage vs. Placebo	Hx of 2 SABs[B] or 2 PTBs[A]	RCT, DB[C] Placebo 1:1	17P: 39 (39 cerclage) Placebo: 40 (40 cerclage)	Booking (12.2 wks av.)	37 wks	17P: 5/31 (16.1%) Placebo: 14/37 (37.8%)	17P: 8/39 Placebo: 3/40
Suvonnakote 1986[5]	17P: 250 mg weekly vs. no treatment	Hx of 1 PTB[A] or 2 late SABs[B]	Non-randomized	17P: 36 No Rx: 39	16 – 20 wks	37 wks	17P: 5/35 (14%) No Rx: 19/39 (49%)	
Hauth 983[6]	17P: 1000 mg weekly vs. Placebo	Active duty military	RCT, DB[C]	17P: 80 Placebo: 88	16 – 20 wks	36 wks	17P: (6.3%) Placebo: (5.7%)	

[A] PTB=Preterm Births

[B] SABs=Spontaneous Abortions

[C] RCT, DB=Randomized Controlled Trial, Double Blind

1 LeVine L. Habitual abortion. A controlled clinical study of progestational therapy. West J Surg Obstet Gynecol. 1964;72:30-6.

2 Papiernik-Berkhauer E. Double blind study of an agent to prevent pre-term delivery among women at increased risk. In: Edition Schering, Serie IV, fiche 3; 65-8; 1970.

3 Johnson JW, Austin KL, Jones GS, Davis GH, King TM. Efficacy of 17 α-hydroxyprogesterone caproate in the prevention of premature labor. N Engl J Med. 1975;293(14): 675-80.

4 Yemini M, Borenstein R, Dreazen E, Apelman Z, Mogilner BM, .Kessler I, et al. Prevention of premature labor by 17 a-hydroxyprogesterone caproate. Am J Obstet Gynecol. 1985;151(5):574-7.

5 Suvonnakote T. Prevention of pre-term labour with progesterone. J Med Assoc Thai.1986; 69(10):538-42.

6 Hauth JC, Oilstrap LC 3rd, Brekken AL, Hauth JM. The effect of 17 α-hydroxyprogesterone caproate on pregnancy outcome in an active-duty military population. Am J Obstet Gynecol. 1983;146(2):187-90.

The study previously conducted that is most comparable to the MFMU Network trial was the double-blind randomized controlled trial conducted by Johnson et al in 1975 at Johns Hopkins University.[26] This study enrolled women with ≥ 2 preterm births, ≥ 2 spontaneous abortions, or a combination of both. Exclusion criteria included: absence of a viable intrauterine pregnancy; failure to enter the study before 24 weeks gestation; and failure to receive a minimum of 3 doses of the assigned medication. Subjects were randomized to receive 17OHP-C 250 mg IM weekly from enrollment into prenatal care until 37 weeks

gestation. Cervical suturing was performed on patients thought to have cervical incompetence (4 in the treatment arm; 3 in the placebo arm). Four patients received isoxsuprine: 2 in the treatment arm; 2 in the placebo arm. Premature birth did not occur in any of the 18 patients receiving 17OHP-C; 9 of 22 patients (41%) receiving placebo had premature birth. The perinatal mortality rate in the 17OHP-C arm was 0% compared to 27% in the placebo arm: of the 7 placebo deaths, 2 were neonatal deaths and 5 were intrauterine deaths.

Other published clinical studies with 17OHP-C have both supported and raised doubt about the effectiveness of 17OHP-C for the prevention of preterm birth. This disparity of opinion prompted the NICHD, via the MFMU Network, to conduct a multicenter placebo-controlled trial to assess the efficacy of 17OHP-C for the prevention of PTB. On June 12, 2003, data from the MFMU Network clinical trial was published in the New England Journal of Medicine, reporting a benefit of 17OHP-C by reducing preterm birth at < 37 weeks.[27] Data from the MFMU Network clinical trial (referred to as Study 17P-CT-002 in this application) provide the primary support for the safety and efficacy of 17OHP-C for the prevention of preterm birth.

2 REGULATORY CONSIDERATIONS AND ISSUES

2.1 Clinical Evidence of Effectiveness

2.1.1 General Considerations

The Division of Reproductive and Urologic Products (hereafter referred to as DRUP or the Division) would typically advise a sponsor developing a drug product for a condition for which there was no previously approved drug product, such as "prevention of preterm birth in pregnant women with a history of at least one spontaneous preterm birth," to conduct 2 adequate and well-controlled clinical trials. The principal reason for such a recommendation is to provide independent substantiation of experimental results. It has been FDA's position that Congress generally intended to require at least 2 adequate and well-controlled studies, each convincing on its own, to establish effectiveness. However, in the 1997 Food and Drug Administration Modernization Act, Congress amended section 505(d) of the Food, Drug, and Cosmetic Act to clarify that the Agency may consider "data from one adequate and well-controlled clinical investigation and confirmatory evidence" to constitute substantial evidence if FDA determines that such data and evidence are sufficient to establish effectiveness.

In NDA 21-945 for Gestiva for prevention of preterm birth, the Applicant has submitted data from only one clinical trial that appears to be adequate and well-controlled (subject to the FDA's inspection of the clinical trial sites and ongoing review of the clinical data). The Division decided to accept this NDA for review in spite of there being only one adequate and well-controlled clinical trial, in part, because of the public health importance of reducing the incidence of preterm birth and its attendant morbidity and mortality and the absence of an approved drug product for this disorder. In addition, there have been examples where the FDA has approved a new drug product based on data from a single adequate and well-controlled clinical trial. In the following sections, the Division provides an overview of the quantity and quality of evidence that is required to approve a new drug product and examples of situations in which data from a single adequate and well-controlled clinical trial has

formed the basis for demonstrating effectiveness. The following discussion is derived from the FDA's Guidance Document entitle *Guidance for Industry, Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products (May 1998).* The complete Guidance can be found in Appendix No. 1 of this background document.

2.1.2 Regulatory Background regarding the Quantity of Evidence Necessary to Support Effectiveness of a Drug Product

In 1962, Congress amended the Federal Food, Drug, and Cosmetic Act to add a requirement that, to obtain marketing approval, manufacturers demonstrate the effectiveness of their products through the conduct of adequate and well-controlled studies. The 1962 Amendments included a provision requiring manufacturers of drug products to establish a drug's effectiveness by "substantial evidence." Substantial evidence was defined in section 505(d) of the Act as "evidence consisting of adequate and well-controlled investigations, including clinical investigations, by experts qualified by scientific training and experience to evaluate the effectiveness of the drug involved, on the basis of which it could fairly and responsibly be concluded by such experts that the drug will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the labeling or proposed labeling thereof."

With regard to quantity, it has been FDA's position that Congress generally intended to require at least 2 adequate and well-controlled studies, each convincing on its own, to establish effectiveness. FDA's position is based on the language in the statute and the legislative history of the 1962 amendments. Language in a Senate report suggested that the phrase "adequate and well-controlled investigations" was designed not only to describe the quality of the required data but the "quantum" of required evidence. Section 505(d) of the Act uses the plural form in defining "substantial evidence" as "adequate and well-controlled investigation**s**, including clinical investigation**s** [underlines added]." Section 505(b) of the Act also uses "investigations" in describing the contents of a new drug application.

Nevertheless, FDA has been flexible within the limits imposed by the congressional scheme, broadly interpreting the statutory requirements to the extent possible where the data on a particular drug were convincing. In some cases, FDA has relied on pertinent information from other adequate and well-controlled studies of a drug, such as studies of other doses and regimens, of other dosage forms, in other stages of disease, in other populations, and of different endpoints, to support a single adequate and well-controlled study demonstrating effectiveness of a new use. In these cases, although there is only one study of the exact new use, there are, in fact, multiple studies supporting the new use, and expert judgment could conclude that the studies together represent substantial evidence of effectiveness. In other cases, FDA has relied on only a single adequate and well-controlled efficacy study to support approval — generally only in cases in which a single multicenter study of excellent design provided highly reliable and statistically strong evidence of an important clinical benefit, such as an effect on survival, and where a confirmatory study would have been difficult to conduct on ethical grounds

2.1.3 Scientific Basis for the Legal Standard

The usual requirement for more than one adequate and well-controlled investigation reflects the need for independent substantiation of experimental results. A single clinical

experimental finding of efficacy, unsupported by other independent evidence, has not usually been considered adequate scientific support for a conclusion of effectiveness. The reasons for this include:

- Any clinical trial may be subject to unanticipated, undetected, systematic biases.
- The inherent variability in biological systems may produce a positive trial result by chance alone. This possibility is acknowledged, and quantified to some extent, in the statistical evaluation of the result of a single efficacy trial. It should be noted, however, that hundreds of randomized clinical efficacy trials are conducted each year with the intent of submitting favorable results to FDA. Even if all drugs tested in such trials were ineffective, one would expect one in forty of those trials to "demonstrate" efficacy by chance alone at conventional levels of statistical significance. Independent substantiation of a favorable result protects against the possibility that a chance occurrence in a single study will lead to an erroneous conclusion that a treatment is effective.
- Results obtained in a single center may be dependent on site or investigator specific factors (e.g., disease definition, concomitant treatment, diet). In such cases, the results, although correct, may not be generalizable to the intended population. This possibility is the primary basis for emphasizing the need for independence in substantiating studies.

Although there are statistical, methodological, and other safeguards to address the identified problems, they are often inadequate to address these problems in a single trial. Independent substantiation of experimental results addresses such problems by providing consistency across more than one study, thus greatly reducing the possibility that a biased, chance, site-specific, or fraudulent result will lead to an erroneous conclusion that a drug is effective.

2.1.4 The Quantity of Evidence to Support Effectiveness

There may be situations in which a single multicenter study, without supporting information from other adequate and well-controlled studies, may provide evidence that a use is effective.

In each of these situations, it is assumed that any studies relied on to support effectiveness meet the requirements for adequate and well-controlled studies as defined in 21 CFR 314.126. It should also be appreciated that reliance on a single study of a given use, whether alone or with substantiation from related trial data, leaves little room for study imperfections or contradictory (nonsupportive) information. In all cases, it is presumed that the single study has been appropriately designed, that the possibility of bias due to baseline imbalance, unblinding, post-hoc changes in analysis, or other factors is judged to be minimal, and that the results reflect a clear prior hypothesis documented in the protocol. Moreover, a single favorable study among several similar attempts that failed to support a finding of effectiveness would not constitute persuasive support for a product use unless there were a strong argument for discounting the outcomes in the studies that failed to show effectiveness.

2.1.5 Evidence of Effectiveness from a Single Study

At present, major clinical efficacy studies are typically multicenter, with clear, prospectively determined clinical and statistical analytic criteria. These studies are less vulnerable to certain biases, are often more generalizable, may achieve very convincing statistical results, and can often be evaluated for internal consistency across subgroups, centers, and multiple endpoints. The added rigor and size of contemporary clinical trials have made it possible to

rely, in certain circumstances, on a single adequate and well-controlled study, without independent substantiation from another controlled trial, as a sufficient scientific and legal basis for approval.

Whether to rely on a single adequate and well-controlled study is inevitably a matter of judgment. A conclusion based on 2 persuasive studies will always be more secure than a conclusion based on a single, comparably persuasive study. For this reason, *reliance on only a single study will generally be limited to situations in which a trial has demonstrated a clinically meaningful effect on mortality, irreversible morbidity, or prevention of a disease with potentially serious outcome and confirmation of the result in a second trial would be logistically impossible or ethically unacceptable.* Repetition of positive trials showing only symptomatic benefit would generally not present the same ethical concerns.

The discussion that follows identifies the characteristics of a single adequate and well-controlled study that could make the study adequate support for an effectiveness claim. Although none of these characteristics is necessarily determinative, the presence of one or more in a study can contribute to a conclusion that the study would be adequate to support an effectiveness claim.

- **Large multicenter study**
  In a large multicenter study in which (1) no single study site provided an unusually large fraction of the subjects and (2) no single investigator or site was disproportionately responsible for the favorable effect seen, the study's internal consistency lessens concerns about lack of generalizability of the finding or an inexplicable result attributable only to the practice of a single investigator. If analysis shows that a single site is largely responsible for the effect, the credibility of a multicenter study is diminished.

- **Consistency across study subsets**
  Frequently, large trials have relatively broad entry criteria and the study populations may be diverse with regard to important covariates such as concomitant or prior therapy, disease stage, age, gender or race. Analysis of the results of such trials for consistency across key patient subsets addresses concerns about generalizability of findings to various populations in a manner that may not be possible with smaller trials or trials with more narrow entry criteria.

- **Multiple endpoints involving different events**
  In some cases, a single study will include several important, prospectively identified primary or secondary endpoints, each of which represents a beneficial, but different, effect. Where a study shows statistically persuasive evidence of an effect on more than one of such endpoints, the internal weight of evidence of the study is enhanced. For example, favorable effects on both death and nonfatal myocardial infarctions in a lipid-lowering, post angioplasty, or post infarction study would, in effect, represent different, but consistent, demonstrations of effectiveness, greatly reducing the possibility that a finding of reduced mortality was a chance occurrence.

  In contrast, a beneficial effect on multiple endpoints that evaluate essentially the same phenomenon and correlate strongly, such as mood change on 2 different depression scales, or SGOT and CPK levels post-infarction, does not significantly enhance the internal weight of the evidence from a single trial.

Although 2 consistent findings within a single study usually provide reassurance that a positive treatment effect is not due to chance, they do not protect against bias in study conduct or biased analyses. For example, a treatment assignment not well balanced for important prognostic variables could lead to an apparent effect on both endpoints. Thus, close scrutiny of study design and conduct are critical to evaluating this type of study.

- **Statistically very persuasive finding**
  In a multicenter study, a very low p-value indicates that the result is highly inconsistent with the null hypothesis of no treatment effect. In some studies it is possible to detect nominally statistically significant results in data from several centers, but, even where that is not possible, an overall extreme result and significance level means that most study centers had similar findings. For example, preventive vaccines for infectious disease indications with a high efficacy rate (e.g., point estimate of efficacy of 80% or higher and a reasonably narrow 95% confidence interval) have been approved based on a single adequate and well-controlled trial.

2.1.6 Reliance on a Single, Multicenter Study — Caveats

While acknowledging the persuasiveness of a single, internally consistent, strong multicenter study, it must be appreciated that even a strong result can represent an isolated or biased result, especially if that study is the only study suggesting efficacy among similar studies. There are examples where the apparent highly favorable effect of drug, studied in what appeared to be a well-designed, placebo-controlled, multicenter trial, resulting in an extreme p-value, has proven to be unrepeatable.

When considering whether to rely on a single multicenter trial, it is critical that the possibility of an incorrect outcome be considered and that all the available data be examined for their potential to either support or undercut reliance on a single multicenter trial.

Inadequacies and inconsistencies in the data, such as lack of pharmacologic rationale and lack of expected other effects accompanying a critical outcome, can weaken the persuasiveness of a single trial. Although an unexplained failure to substantiate the results of a favorable study in a second controlled trial is not proof that the favorable study was in error — studies of effective agents can fail to show efficacy for a variety of reasons — it is often a reason not to rely on the single favorable study.

2.1.7 Documentation of the Quality of Evidence Supporting an Effectiveness Claim

When submitting the requisite quantity of data to support approval of a new product or new use of an approved product, sponsors must also document that the studies were adequately designed and conducted. To demonstrate that a trial supporting an effectiveness claim is adequate and well-controlled, extensive documentation of trial planning, protocols, conduct, and data handling is usually submitted to the Agency, and detailed subject records are made available at the clinical sites.

From a scientific standpoint, however, it is recognized that the extent of documentation necessary depends on the particular study, the types of data involved, and the other evidence available to support the claim. Therefore, the Agency is able to accept different levels of documentation of data quality, as long as the adequacy of the scientific evidence can be assured. The issues of prime importance in documenting the quality of the evidence are

(1) the completeness of the documentation and (2) the ability to access the primary study data and the original study-related records (e.g., subjects' medical records, drug accountability records) for the purposes of verifying the data submitted as evidence.

In practice, to achieve a high level of documentation, studies supporting claims are ordinarily conducted in accordance with good clinical practices (GCPs). Sponsors routinely monitor all clinical sites, and FDA routinely has access to the original clinical protocols, primary data, clinical site source documents for on-site audits, and complete study reports.

However, situations often arise in which studies that evaluate the efficacy of a drug product lack the full documentation described above (for example, full subject records may not be available) or in which the study was conducted with less monitoring than is ordinarily seen in commercially sponsored trials. Under certain circumstances, it is possible for sponsors to rely on such studies to support effectiveness claims, despite less than usual documentation or monitoring. Some of those circumstances are described below.

Reliance on Studies with Alternative, Less Intensive Quality Control/On-Site Monitoring

Industry-sponsored studies typically use extensive on-site and central monitoring and auditing procedures to assure data quality. Studies supported by other sponsors may employ less stringent procedures and may use no on-site monitoring at all. An International Conference on Harmonisation guideline on good clinical practices (*"International Conference on Harmonisation Guidance for Industry E6, Good Clinical Practice: Consolidated Guideline, April 1996"*) emphasizes that the extent of monitoring in a trial should be based on trial-specific factors (e.g., design, complexity, size, and type of study outcome measures) and that different degrees of on-site monitoring can be appropriate. In recent years, many credible and valuable studies conducted by government or independent study groups, often with important mortality outcomes, had very little on-site monitoring. These studies have addressed quality control in other ways, such as by close control and review of documentation and extensive guidance and planning efforts with investigators.

2.2 Discussions between Adeza and the Division

After data from Study 17P-CT-002 were published in the New England Journal of Medicine (Meis et al. 2003),[27] Adeza met with the Division to discuss the possibility of submitting an NDA for 17OHP-C for prevention of preterm birth.

The Division conveyed several recommendations and concerns to the Applicant during this and subsequent meetings. These included the following:

- A major concern was the lack of follow-up data, beyond the period of initial hospital assessment, of babies in which the mother received 17OHP-C for the prevention of preterm birth. The Division requested that the applicant obtain follow-up data on infants through at least 2 years of age.
- A second major concern related to the drug product(s) used during the trial. The Sponsor was informed that complete chemistry, manufacturing and control (CMC) information would need to be provided about the drug product, including its purity and potency. The applicant would need to provide information that the drug product used in the NIH sponsored clinical trial and the to-be-marketed formulation would be comparable.

- The Division had some concerns about outcomes of Study 17P-CT-002 and the adequacy of these outcomes to support approval of a new drug product for marketing in the U.S, particularly since the NDA supporting the safety and effectiveness of 17OHP-C would be based primarily on the outcome of a single clinical trial. These concerns included:
  - The lack of any suggestion of improvement in overall mortality in the 17OHP-C treated subjects compared to the placebo treated subjects.
  - Clinical Trial 17P-CT-002 did not show a statistically robust effect for reducing the number of births at gestational ages <32 weeks, when infant morbidity/mortality is a much greater problem in the U.S. The Division, however, recognized that the trial was not powered for this endpoint.
  - The primary endpoint of Clinical Trial 17P-CY-002 was a surrogate for pregnancy outcome (neonatal/infant morbidity and mortality). The Division indicated that its review would focus on what it believed to be the most important outcomes (overall survival of fetuses/infants and a significant reduction in serious morbidities from the time of enrollment rather than merely an increase in gestational age, without other accompanying clinical benefits).
  - Normally, either 2 adequate and well-controlled studies or a single study with a robust and compelling outcome and strong supporting data would be required to support approval of a new drug product. There was a possibility that the data from Trial 17P-CT-002 would not be sufficient to demonstrate that 17OHP-C is safe and effective for the prevention of preterm birth.

3 OVERVIEW OF CLINICAL DATA IN NDA 21-945

In support of their application for the use of 17OHP-C for the prevention of preterm birth the Applicant submitted data from 2 active treatment clinical trials and a follow-up safety study: Study 17P-IF -001; Study 17P-CT-002 and follow up study 17P-FU. An overview of these studies is presented in Table 2.

Table 2 Studies of 17OHP-C for Prevention of Recurrent Preterm Births

Protocol # /Status	Study Design	Study Population	Treatment Dose	Duration of Drug Treatment	Number of Subjects	Race: Black/Non-Black	Mean Age (Range)
17P-IF-001 Terminated [A] Mar 1999	Double-blind, Placebo-controlled, Randomized 2:1 active treatment to Placebo	Pregnant women with previous spontaneous preterm birth	250 mg/week	Weekly injections beginning from 16^0 to 20^6 wks gestation until 37^0 wks gestation or delivery	Total: 150 17P: 94 Placebo: 56	Total: 95/55 17P: 54/40 Placebo: 41/15	26.2 yr (17, 42)
17P-CT-002 Completed [B] Aug 2002	Double-blind, Placebo-controlled, Randomized 2:1 active treatment to Placebo	Pregnant women with previous spontaneous preterm birth	250 mg/week	Weekly injections beginning from 16^0 to 20^6 wks gestation until 37^0 wks gestation or delivery	Total: 463 17P: 310 Placebo: 153	Total: 273/190 17P: 183/127 Placebo: 90/63	26.2 yr (16, 43)
17P-FU Completed Nov 2005	Observational long-term safety follow-up for Study 17P-CT-002	Infants discharged live in Study 17P-CT-002	None	No study treatment was administered	Total: 278 17P: 194 Placebo: 84	Total: 152/126 17P: 105/89 Placebo: 47/37	47.4 mo (30, 64)

[A] Study 17P-IF-001 was terminated early by the Sponsor when the manufacturer recalled the study drug. The last subject visit was in August 1999. Of the 150 subjects, only 60.6% (57/94) of subjects randomized to 17OHP-C and 51.8% (29/56) of subjects randomized to placebo completed study treatment to 36^6 weeks of gestation or delivery.

[B] An independent Data and Safety Monitoring Committee (DSMC) reviewed the study data after 400 subjects had completed the study. Based on that interim dataset, the primary endpoint, birth $<37^0$ weeks of gestation, was significantly reduced and the p-value was below the p-value specified in predefined stopping rules. The DSMC recommended that enrollment in the study be stopped, so that no new subjects would be assigned placebo. By the time the study was stopped, 463 subjects had been enrolled, which was 92.5% of the proposed sample size of 500 subjects.

Initial Formulation Study (Study 17P-IF-001)

This study began in February 1998, but treatment was terminated in March 1999 because the active study drug (17OHP-C) was recalled by its manufacturer, under the direction of the FDA, due to violations of manufacturing practices potentially affecting the potency of the drug. At the time of termination, only 150 of the proposed 500 subjects had been randomized, and no data analysis had been done. Ninety subjects completed the treatment regimen before the study was stopped: 57 (61%) of the 17OHP-C subjects and 29 (52%) of the placebo subjects. The study drug used in this terminated study is referred to as the Initial Formulation (IF). The data collected from subjects enrolled in the terminated study were analyzed separately in the NDA and the results are also summarized separately.

Principal Clinical Trial (Study 17P-CT-002)

This study, which began in October 1999, randomized 463 subjects who had at least one documented prior spontaneous preterm birth of a singleton, non-anomalous fetus. Of these, 418 subjects (90.3%) completed dosing through 36^6 weeks or birth: 279 (90.0%) in the

17OHP-C group and 139 (90.8%) in the placebo group. This study was terminated prior to enrolling the proposed 500 subjects because the prespecified stopping criterion for efficacy was attained at an interim analysis.

Follow-up of Children from the 17P-CT-002 trial (Study 17P-FU)
This was a follow-up to Study 17P-CT-002. The follow-up study collected data with a validated child development instrument, the Ages and Stages Questionnaire (ASQ), a Survey Questionnaire concerning the health and development of the child, and a physical examination. The children were at least 2 years of age at the time of the follow-up assessments. The primary objective of this study was to determine whether there was a difference in achievement of developmental milestones and physical health between children born to women who received weekly intramuscular injections of 17OHP-C compared with placebo during the pregnancy in Study 17P-CT-002.

4 PRIMARY EFFICACY AND SAFETY CLINICAL TRIAL

Study 17P-CT-002: "A Randomized Trial of 17α-Hydroxyprogesterone Caproate for Prevention of Preterm Birth in High Risk Women"

4.1 Background Information

The National Institute of Child Health and Human Development (NICHD) created the Maternal-Fetal Medicine Units (MFMU) Network in 1986 to focus on clinical questions in maternal fetal medicine and obstetrics, particularly with respect to the continuing problem of preterm birth. Operating under cooperative agreements at the time this study was conducted, the MFMU Network comprised 19 university-based clinical centers and a data-coordinating center, the Biostatistical Coordinating Center (BCC) at George Washington University. The NICHD/MFMU Network was responsible for operational issues including site monitoring and project management for this study

The plan was to conduct one multicenter, randomized, placebo-controlled, double-blinded study on the efficacy and safety of 17OHP-C in pregnant women at high risk for preterm birth. Study 17P-IF-001 enrolled its first subject in February 1998, but had to be terminated early in March 1999 after only one-third of the proposed subjects were enrolled. None of the data had been analyzed at the time of termination. This termination occurred because the study drug (17OHP-C) was recalled by its manufacturer at the request of the FDA as described in Section 3.

The clinical trial was started afresh in October 1999 using study drug from a new manufacturer and is referred as Study 17P-CT-002. The data collected from subjects enrolled in the terminated Study 17P-IF-001 were not merged with data collected in Study 17P-CT-002 nor were they provided in the Report for Study 17P-CT-002.

4.2 Study Drugs

Active study drug consisted of 17α-hydroxyprogesterone caproate (250 mg/mL) in castor oil with 46% benzyl benzoate and 2% benzyl alcohol. Inactive (placebo) study drug was identical to the active drug product but did not contain 17OHP-C. Study drugs were administered once weekly by intramuscular injection.

4.3 Overview of Protocol for Study 17P-CT-002

Study 17P-CT-002 was conducted at 19 investigational sites in the United States. All principal investigators were members of the NICHD MFMU Network. Certification of each study center was required before recruitment of subjects.

The study was a randomized, placebo-controlled, efficacy and safety study of 17OHP-C in pregnant women, from 16^0 to 20^6 weeks gestation, who had a history of spontaneous preterm birth, defined as delivery from 20^0 to 36^6 weeks gestation following spontaneous preterm labor (PTL) or preterm premature rupture of membranes (pPROM). The requirement that the gestational age be at least 16^0 weeks and no more than 20^6 weeks was instituted in order to initiate treatment after the first trimester, but before the gestational age at which a preterm birth, by definition, could occur.

Prior to randomization into the clinical trial, an injection of the placebo drug product was administered to potential subjects from 15^0 to 20^3 weeks gestation, to assess the subject's tolerability to the injection. Qualifying subjects were randomized in a 2:1 ratio to 17OHP-C or placebo. Study drug was administered weekly by intramuscular injection through 36^6 weeks gestation or delivery, whichever occurred first.

4.3.1 Inclusion/Exclusion Criteria

Inclusion Criteria. Subjects had to meet all of the following criteria at screening to be eligible for enrollment into the study:

1. Gestational age between 16^0 weeks and 20^6 weeks at the time of randomization, based on clinical information and evaluation of the first ultrasound.
2. Documented history of a previous singleton spontaneous preterm birth. Spontaneous preterm birth was defined as delivery from 20^0 to 36^6 weeks gestation following spontaneous preterm labor or preterm premature rupture of membranes. Where possible, the gestational age of the previous preterm birth (referred to as the qualifying birth) was determined. If the gestational age at delivery was obtained directly from the medical record and more than one gestational age appeared, the latest was used. The qualifying delivery could not be an antepartum stillbirth.

Exclusion Criteria. If any of the following criteria applied, the subject was not eligible to enroll into the study:

1. Multifetal gestation.
2. Known major fetal anomaly or fetal demise. An ultrasound examination after 14 weeks gestation had to be performed to rule out fetal anomalies.
3. Progesterone treatment during current pregnancy.
4. Heparin therapy during current pregnancy or history of thromboembolic disease.
5. Maternal medical/obstetrical complications including:
    a. Current or planned cerclage;
    b. Hypertension requiring medication;
    c. Seizure disorder.

6. Prenatal follow-up or delivery planned elsewhere (unless the study visits could be made as scheduled and complete outcome information obtained).
7. A 14^0 to 20^6 week ultrasound could not be arranged before randomization.
8. Participation in an antenatal study in which the clinical status or intervention could have influenced gestational age at delivery. Subjects enrolled in any of the following MFMU Network studies during this period were ineligible for the trial: "Randomized Clinical Trials of the Effect of Metronidazole on Pregnancy Outcome in Women Infected with T. Vaginalis or Bacterial Vaginosis," "Randomized Trial of Metronidazole Plus Erythromycin to Prevent Preterm Birth in Women with Elevated Cervical/Vaginal Oncofetal Fibronectin," "Randomized Clinical Trial of Theophylline versus Inhaled Beclomethasone," and "The Effects of Asthma and Treatment Regimens on Perinatal Outcome."
9. Participation in this trial in a previous pregnancy. Subjects who were screened in a previous pregnancy, but not randomized, were not excluded.

4.3.2 Endpoints

Primary Objective. The primary per protocol objective of this study was to determine if, compared with placebo, 17OHP-C treatment initiated before 21^0 weeks gestation reduces the risk of preterm birth ($<37^0$ weeks gestation) in women who have previously experienced a spontaneous preterm birth.

All deliveries occurring from the time of randomization through 36^6 weeks gestation, including miscarriages (i.e., spontaneous abortions) and elective abortions, were counted in the primary outcome.

Secondary Objectives. The secondary objectives defined in the protocol were to determine the following in women with a previous spontaneous preterm birth:

- If treatment with 17OHP-C reduces the use of tocolytic therapy and/or cervical cerclage.
- If treatment with 17OHP-C reduces neonatal morbidity/mortality.

Neonatal outcomes considered secondary efficacy measures included: birthweight; score reflecting condition of neonate (Apgar score); admission to the neonatal intensive care unit (NICU); infant hospital days; number of days of neonatal respiratory therapy; stillbirths; neonatal deaths; neonates with respiratory distress syndrome (RDS); intraventricular hemorrhage (IVH); bronchopulmonary dysplasia (BPD); necrotizing enterocolitis (NEC); early onset of neonatal sepsis; seizures; retinopathy of prematurity; and transient tachypnea. In addition, the percentage of infants who received ventilator support, and the percentage of infants who received supplemental oxygen were provided.

Based on communications with the FDA, the following secondary endpoints were added to the analyses:

- If treatment with 17OHP-C, compared to placebo, reduces the risk of preterm birth of $<35^0$ weeks gestations.
- If treatment with 17OHP-C, compared to placebo, reduces the risk of preterm birth of $<32^0$ weeks gestations.

- If treatment with 17OHP-C, compared to placebo, reduces overall neonatal morbidity based on a composite measure of neonatal morbidity.

4.3.3 Statistical Methods/Sample Size Determination

Applicant's Analyses. All statistical comparisons were between 17OHP-C and placebo. Except where explicitly indicated, data were pooled across study centers for all statistical analyses. Subjects were analyzed based on the group to which they were randomized.

Summary statistics consisted of numbers and percentages of subjects for categorical measures and were compared for statistical significance between treatment groups using the chi-square test, Fisher's Exact test, or the Wilcoxon Rank Sum test for ordered categorical data. For categorical variables, percentages were calculated based on available data.

Summary statistics consisted of means, medians, standard deviations, and minimum and maximum values for continuous measures and were compared for statistical significance between the treatment groups using the Wilcoxon Rank Sum test.

All statistical tests were reported as 2-sided p-values. The final primary efficacy analysis utilized the Type 1 α=0.034 level of statistical significance as required by the O'Brien Fleming boundary. For all other analyses, no adjustments were made for multiple comparisons and a nominal α=0.05 level of statistical significance was used.

4.4 Demographics, Concomitant Medication Use, and Subject Disposition

4.4.1 Demographics and Obstetrical History

The subjects randomized to the 2 treatment groups (17OHP-C vs. placebo, respectively) were comparable in mean age, race or ethnic group, mean BMI prior to pregnancy, marital status, mean years of education, and substance use during pregnancy. The mean age of the subjects was 26.2 years (26.0 vs. 26.0 years) and their mean pre-pregnancy BMI was 26.6 kg/m^2 (26.9 vs. 26.0 kg/m^2). Half of the subjects were married or living with a partner (51% vs. 46%), while 39.5% had never been married (38% vs. 42%). More than half of the subjects were African American (59% in each group); and 4% had a history of diabetes (4% vs. 3%). During the study pregnancy but prior to randomization, 22% had smoked (23% vs. 20%), 8% had consumed alcoholic drinks (9% vs. 6%), and 3% had used street drugs (4% vs. 3%).

Obstetrical histories were comparable in the 17OHP-C and placebo groups for gestational age at randomization (18.9 vs. 18.8 weeks), gestational age of qualifying delivery (30.6 and 31.3 weeks), number of previous term deliveries (0.8 and 0.7); percentage with previous miscarriages (30.0% vs. 37.3%) and stillbirths (10.0% vs. 8.5%). (See Table 3.)

Division's Comment

- *The 17OHP-C subjects had statistically significantly fewer previous preterm births (1.4 vs. 1.6), fewer previous SPTB (1.3 vs. 1.5), and a lower percentage of subjects with >1 previous preterm birth (27.7% vs. 41.2%). They may therefore represent a lower-risk group as compared to the placebo subjects.*

One-third of the subjects in each treatment group had an infection during the study pregnancy prior to randomization (32% in 17OHP-C vs. 36% in placebo groups). The types of infections prior to randomization were similar across the treatment groups. The most

common infections were bacterial vaginosis (13% in both treatment groups), urinary tract infections (12% vs. 13%), and Chlamydia infections (3.9% vs. 4.6%).

A smaller percentage of subjects randomized to 17OHP-C used corticosteroids during the study pregnancy prior to randomization (1.6% vs. 5.2%); the difference was due to a lower use of inhaled corticosteroids in the 17OHP-C group (0.3% vs. 4.6%).

Table 3 Obstetrical History

Obstetrical History	17OHP-C (N=310)	Placebo (N=153)	P-value[A]
Gestational age of qualifying birth, wk			
Mean (SD)	30.6 (4.6)	31.3 (4.2)	
Min, Max	20, 36	20, 36	
No. of previous preterm births (PTBs)			
Mean (SD)	1.4 (0.7)	1.6 (0.9)	<0.05
Min, Max	1, 5	1, 6	
>1 Previous preterm birth, n (%)	86 (27.7)	63 (41.2)	<0.05
No. of previous spontaneous PTBs			
Mean (SD)	1.3 (0.7)	1.5 (0.9)	<0.05
Min, Max	1, 5	1, 6	
No. of previous term deliveries			
Mean (SD)	0.8 (1.1)	0.7 (1.0)	
Min, Max	0, 7	0, 5	
Previous miscarriage, n (%)	93 (30.0)	57 (37.3)	
Previous stillbirth, n (%)	31 (10.0)	13 (8.5)	
Infection during pregnancy (before randomization), n (%)	98 (31.6)	55 (35.9)	
Corticosteroids during pregnancy (before randomization), n (%)	5 (1.6)	8 (5.2)	<0.05
Duration of gestation at randomization, wk			
Mean (SD)	18.9 (1.4)	18.8 (1.5)	
Min, Max	16, 21	16, 21	

[A] Only p-values ≤ 0.05 shown.

Source: Table 11-2, Final Report for Study 17-CT-002.

4.4.2 Concomitant Medication Use

No attempt was made to mandate clinical management of the subjects during the study. The percentages of subjects who received any type of corticosteroids (16.8% vs. 19.6%), antibiotic therapy (31.6% vs. 23.5%), or tocolytic therapy (12.9% vs. 11.8%) were not significantly different between the 17OHP-C and placebo groups. The most common (>5% of subjects) type of corticosteroid used after randomization was parenteral corticosteroids (14.2% in the 17OHP-C group vs. 13.7% in the placebo group). The most common types of antibiotics were penicillin (17.7% vs. 14.4%), oral metronidazole (10.3% vs. 5.2%), and erythromycin (8.7% vs. 8.5%).

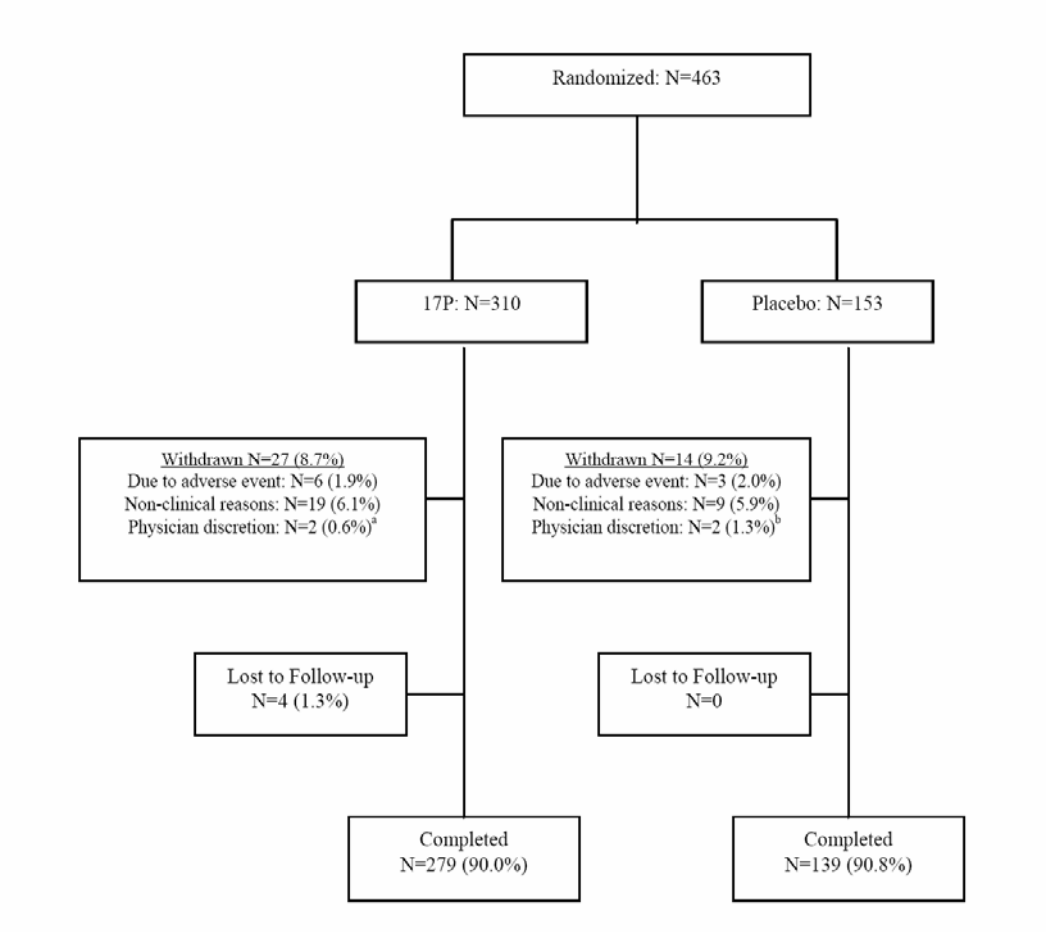
The percentage of subjects using the following concomitant medications differed between the 17OHP-C and placebo groups: inhaled corticosteroids (1.9% vs. 4.6%), oral metronidazole

(10.3% vs. 5.2%), and nitrofurantoin (4.2% vs. 1.3%). Oral metronidazole was administered for bacterial vaginosis or Trichomonas vaginalis and nitrofurantoin was administered for urinary tract infections, which suggests that a slightly higher rate of these infections occurred in the 17OHP-C group during the study pregnancy.

4.4.3 Subject Disposition

A total of 463 subjects were randomized at 19 study centers in the U.S (Figure 1). Four hundred eighteen (418; 90.3%) subjects completed injections through 36^{6} weeks gestation or delivery, whichever occurred first: 279 (90.0%) in the 17OHP-C group and 139 (90.8%) in the placebo group. Early discontinuation of treatment with study drug occurred at a similar rate in both treatment groups (8.7% 17OHP-C vs. 9.2% placebo). Most of these subjects discontinued due to "non-clinical reasons," which were not further defined by the Applicant (6.1% vs. 5.9%); those potentially due to adverse events (AEs) are discussed in Section 4.6.6. Four (<1.0%) subjects, all in the 17OHP-C group, were lost to follow-up.

Figure 1 Overview of Subject Disposition in Study 17P-CT-002



Note: "Withdrawn from the study" was defined as the patient no longer received study drug. "Lost to follow-up" was defined as the patient's delivery data could not be obtained. "Completed the study" was defined as the patient did not withdraw from the study and was not lost to follow-up.

[a] In the 17P group, Investigators stopped the participation of one patient due to injection site reactions and another patient due to pPROM, which was not considered an AE. Therefore, 7 (2.2%) patients in the 17P group discontinued due to AEs.

[b] In the placebo group, Investigators stopped the participation of one patient due to a potential allergic reaction and another patient due to pPROM, which was not considered an AE. Therefore, 4 (2.6%) patients in the placebo group discontinued due to AEs.

Source: Section 10.1, Figure 10-1, Final Report for Study 17-CT-002.

4.5 Efficacy Outcomes

4.5.1 Primary Endpoint (Applicant's Analyses)

The proportions of deliveries prior to 37^0 weeks gestation based on the ITT population and on all available data are summarized in Table 4. In the ITT population, 115 of 310 ((37.1%) had a delivery prior to 37^0 weeks gestation. In the placebo group, 84 of 153 subjects (54.9%) had a delivery prior to 37^0 weeks gestation. The difference was statistically significant.

Table 4 Percentages of Subjects with Delivery <37^0 Weeks Gestation (Sponsor's Analysis)

Data Source	17P		Placebo		Nominal P-value [A]	Treatment difference and 95% Confidence Interval [B]
	N	n (%)	N	n (%)		
ITT population	310	115 (37.1)	153	84 (54.9)	0.0003	-17.8% [-28%, -7%]
Only available data	306	111 (36.3)	153	84 (54.9)	0.0000	-18.6% [-29%, -8%]

ITT population was all randomized subjects. The 4 subjects with missing outcome data were classified as having a preterm birth of <37^0 weeks (i.e., treatment failure). "Only available data" does not include the 4 subjects with missing outcome data.

[A] Chi-square test. Adjusting for interim analyses, p-values should be compared to 0.035 rather than the usual 0.05.

[B] Confidence interval (CI) calculated by FDA, adjusted for the 2 interim analyses and the final analysis. To preserve the overall Type I error rate of 0.05, a p-value boundary of 0.035 was used for the adjustment (equivalent to a 96.5% confidence interval).

Source: Modified from Table 11-3, Final Report for Study 17P-CT-002.

Subjects who delivered prior to 37^0 weeks gestation also were classified (1) by the gestational age of the previous qualifying SPTB using the intervals of 20^0-<28^0 weeks, 28^0-<32^0 weeks, 32^0-<35^0 weeks, and 35^0-<37^0 weeks), (2) by race (African American [non-Hispanic Black] and Non-Black), and (3) by number of previous preterm births (1, 2, and ≥3) (see Table 5)

Table 5 Percentages of Subjects with Delivery <37^0 Weeks by Gestational Age of Qualifying Birth, Race, and Number of Previous Preterm Deliveries

Characteristic	17OHP-C n/N (%)	Placebo n/N (%)
Previous SPTB (qualifying birth) by gestational age		
20^0 - <28^0 weeks	33/82 (40.2)	19/29 (65.5)
28^0 - <32^0 weeks	21/66 (31.8)	17/30 (56.7)
32^0 - <35^0 weeks	30/84 (35.7)	27/55 (49.1)
35^0 - <37^0 weeks	31/78 (39.7)	21/39 (53.8)
Race		
Black	66/183 (36.1)	47/90 (52.2)
Non-Black	49/127 (38.6)	37/63 (58.7)
Number of previous preterm births (PTBs)		
1 prior PTB	74/224 (33.0)	40/90 (44.4)
2 prior PTB	27/56 (48.2)	31/46 (67.4)
≥3 prior PTB	14/30 (46.7)	13/17 (76.5)

Data based on ITT Population (all randomized subjects). The 4 subjects with missing outcome data were classified as having a preterm birth <37^0 weeks (i.e., treatment failure).

Abbreviations:

SPTB = spontaneous preterm birth; PTB = preterm birth.

n = number of subjects in a specific category who delivered study pregnancy at <37^0 weeks gestation
N = total number of subjects overall in a specific category.

Source: Table 11-4, Final Report for Study 17-CT-002.

Rates of preterm birth at $<37^0$ weeks did not appear to differ significantly according to the gestational age of the qualifying delivery in either treatment group (with the possible exception of the category of 20^0 - $<28^0$ weeks in the placebo group). For all intervals of gestational age, the rates of preterm birth $<37^0$ weeks were numerically lower in the 17OHP-C treatment group.

The percentage of Black subjects in Study 17P-CT-002 was 59% in both groups. 17OHP-C reduced the rate of preterm birth of $<37^0$ weeks gestation compared to placebo for both the Black (36.1% vs. 52.2%) and the Non-Black (38.6% vs. 58.7%) populations.

Subjects with more than one previous preterm birth, regardless of treatment group, had numerically increased rates of preterm births for the study pregnancy compared to subjects with only one previous preterm birth. The rates of preterm births in the 17OHP-C treatment group, compared with placebo, were numerically lower for subjects with one previous preterm birth (33% vs. 44%), 2 previous preterm births (48% vs. 67%), and 3 or more previous preterm births (47% vs. 77%). If the last 2 categories were combined, the incidence of preterm birth in this study for subjects with >1 previous preterm birth was 48% in the 17OHP-C group compared with 70% in the placebo group.

Division's Comment

- *Treatment with 17OHP-C reduces preterm births < 37 weeks gestation.*
- *The reduction in preterm birth appeared independent of race, number of qualifying preterm deliveries, and gestational age of qualifying preterm birth.*

4.5.2 Secondary Endpoints

4.5.2.1 Proportion of Deliveries <35 and <32 Weeks Gestational Age (Applicant's Analysis)

At the request of the Division, the Applicant also calculated the proportion of deliveries $<35^0$ weeks gestation and $<32^0$ weeks gestation because of the increasing morbidity associated with earlier premature deliveries. The proportion of deliveries $<35^0$ weeks gestation (21.6% vs. 30.7%) and $<32^0$ weeks gestation (12.6% vs. 19.6%) were lower in the 17OHP-C group compared with the placebo group (see Table 6).

Table 6 Percentages of Subjects with Delivery $<35^0$ and $<32^0$ Weeks Gestation (Applicant's Analysis)

Pregnancy Outcome	17P N=310 n (%)	Placebo N=153 n (%)	Nominal P-value [A]
Delivery $<35^0$	67 (21.6)	47 (30.7)	0.0324
Delivery $<32^0$	39 (12.6)	30 (19.6)	0.0458

Data presented are from the ITT population (i.e., all randomized subjects). The 4 subjects with missing outcome data were classified as having a preterm birth $<37^0$ weeks (i.e., treatment failure).

[A] Adjusting for interim analyses, p-values should be compared to 0.035 rather than the usual 0.05.

Source: Table 11-5, Final Report for Study 17-CT-002.

Division's Comments

- *The p-values presented in Table 6 should be interpreted with caution for several reasons: (1) there were 2 interim analyses and a final analysis and (2) multiple endpoints, likely to be correlated with each other and with the primary endpoint, were analyzed. The adjustment to the p-value that should be used for analyses of multiple endpoints in this setting is not clear. To declare statistical significance, the p-value boundary is likely smaller than the 0.035 used for analysis of the primary endpoint.*
- *Thus, the difference in deliveries at $<35^0$ weeks may be suggestive of a treatment effect but not statistically significant.*

4.5.2.2 Proportion of Deliveries <35 and <32 Weeks Gestational Age (Division's Analysis)

The Division's analysis of the effects of treatment with 17OHP-C, as compared to placebo, on the percentage of deliveries at $<37^0$, $<35^0$, $<32^0$, and $<28^0$ weeks gestation is shown in Table 7. At each of weeks $<37^0$, $<35^0$, and $<32^0$, the percentage of deliveries was numerically lower in the 17OHP-C treatment arm. The point estimates of the differences between the percentage of births at each gestational age ranged from -17.8% (at $<37^0$) to -7.0% (at $<32^0$). However, the upper limits of the 95% confidence intervals (adjusted to preserve the overall Type I error rate of 0.05) of the differences between treatment groups suggest that the true rate of preterm deliveries could be as much as 0.3% and 0.8% higher in the 17OHP-C groups at $<35^0$ weeks and $<32^0$ weeks gestation, respectively.

There was no difference between treatment groups for the percentages of deliveries $<28^0$ weeks.

Table 7 Percentages of Subjects with Delivery $<37^0$, $<35^0$, $<32^0$, and $<28^0$ Weeks Gestation (ITT Population, Division's Analysis)

Time of Delivery (Gestational Age)	17OHP-C (N=310)	Placebo (N=153)	Treatment difference [A] and 95% Confidence Interval [B]
	%	%	
$<37^0$ weeks	37.1	54.9	-17.8% [-28%, -7.0%]
$<35^0$ weeks	21.6	30.7	-9.1% [-18%, 0.3%]
$<32^0$ weeks	12.6	19.6	-7.0% [-14%, 0.8%]
$<28^0$ weeks	10.0	10.5	-0.5% [-6.9%, 5.9%]

[A] Chi-square test.

[B] The confidence intervals, based on a t-test, are adjusted for the 2 interim analyses and the final analysis. To preserve the overall Type I error rate of 0.05, the final p-value boundary of 0.035 was used for the adjustment (equivalent to a 96.5% confidence interval).

Source: FDA statistical analysis of Applicant's data from Study 17P-CT-002.

Division's Comment

- *The 95% confidence intervals for the difference between treatment groups for deliveries $<37^0$ weeks gestation suggest that the true rate of preterm deliveries in the 17OHP-C group could range from 7 to 28% lower than the rate in the placebo group. This finding*

supports the Applicant's claim that treatment with 17OHP-C, compared to placebo, had a statistically significantly effect in reducing the proportion of deliveries <37^0 weeks.

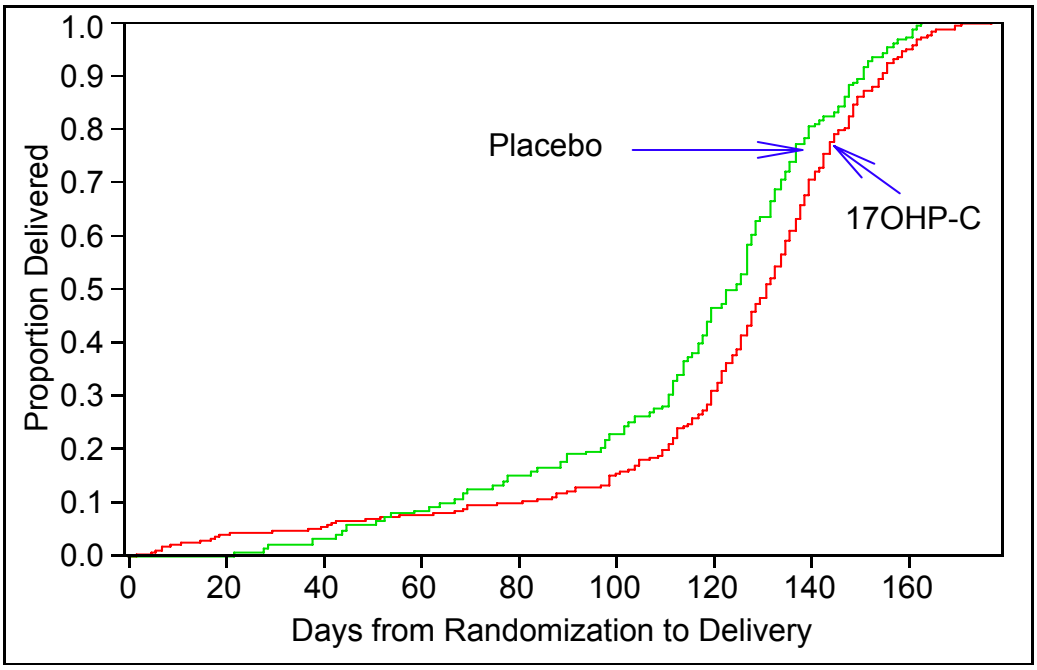
- *The upper limits of the 95% confidence intervals for the differences between treatment groups for deliveries at <35^0 weeks and <32^0 weeks gestation suggest the true rate of preterm deliveries in the 17OHP-C group could be as much as 0.3% and 0.8% higher, respectively, than that in the placebo group. This finding does not allow a conclusion as to whether there is a difference in the true rate of preterm delivery between the treatment groups at <35^0 weeks and <32^0 weeks gestation. If further adjustment of the 95% confidence interval were required (see Division's comment in Section 4.5.2.1), there would be greater doubt as to whether this clinical trial had demonstrated a true difference in the rates of deliveries between the treatment groups at <35^0 weeks and <32^0 weeks gestation.*
- *The Division recognizes that this clinical trial was not powered to demonstrate a difference in the rates of deliveries between the 2 treatment groups at either <35^0 weeks or <32^0 weeks gestation. However, because the Applicant is seeking approval for 17OHP-C based on (1) only a single clinical trial and (2) a surrogate endpoint of neonatal/infant morbidity and mortality, inability to demonstrate a robust effect at either <35^0 weeks or <32^0 weeks gestation is an important consideration in assessing the overall effectiveness of 17OHP-C for the proposed indication.*

4.5.2.3 Mean Gestational Age at Delivery and Duration of Pregnancies

The mean gestational age at delivery for subjects with available outcome data (306 in the 17OHP-C group and 153 in the placebo group) was one week greater in the 17OHP-C group (36.2 weeks vs. 35.2 weeks). The gestational ages at delivery ranged from 18.1 to 41.6 weeks. The median prolongation of pregnancy (defined as the time from randomization until delivery or date that the subject was last confirmed to be pregnant) was higher in the 17OHP-C group compared to the placebo group (131 days vs. 125 days).

A plot of the proportion of subjects delivered as a function of time (days) after randomization is provided in Figure 2. During the period from randomization through approximately 6-7 weeks post-randomization, the proportion of subjects who had delivered was numerically greater in the 17OHP-C treatment group. Thereafter, the proportion of subjects who had delivered was numerically greater in the placebo treatment group at all times through at least Day 150 post-randomization.

Figure 2 Proportion of Subjects Delivered after Onset of Treatment (Study 17P-CT-002)



Source: FDA statistical analysis of Applicant's data from Study 17P-CT-002.

Division's Comment

- *The increased proportion of delivered subjects in the 17OHP-C group, relative to the placebo group, during the first 6 weeks after randomization was due in part to the 5 miscarriages (spontaneous abortions) at <20 weeks gestation in the 17OHP-C group. No miscarriages (spontaneous abortions) at <20 weeks gestation were reported in the placebo group. Whether treatment with 17OHP-C contributed to these early pregnancy losses is not known.*
- *A second randomized clinical trial (or data from other sources) would be helpful in assessing whether treatment with 17OHP-C may be associated with an increase in early pregnancy loss at <20 weeks gestation.*

4.5.2.4 Other Secondary Efficacy Endpoints

The percentage of subjects who were given tocolytic agents during the study was similar in the 2 treatment groups (12.9% vs. 11.8%). The incidence of cerclage placement was also similar in both treatment groups (1.6% vs. 1.3%).

The incidence of caesarian section (C-section) in the 17OHP-C group was similar to that in the placebo group (25.2% vs. 26.8%). The most common reasons for a C-section in the 17OHP-C and placebo groups, respectively, were previous C-section (44.2% vs. 41.5%), abnormal presentation (23.4% vs. 29.3%), and fetal distress (14.3% vs. 19.5%).

4.5.3 Miscarriages, Stillbirths, and Neonatal Deaths

The incidences of miscarriages and stillbirths are summarized in Table 8 and discussed in more detail in Section 4.6.2. Five (1.6%) subjects, all in the 17OHP-C group, experienced miscarriages. No subject in the placebo group miscarried.

The incidence of stillbirths was slightly higher in the 17OHP-C group, but the difference was not statistically significant. Eight subjects had stillbirths: 6 (2.0%) subjects in the 17OHP-C group and 2 (1.3%) subjects in the placebo group. Six of the 8 stillbirths were antepartum stillbirths (fetal deaths in utero) and 2 occurred intrapartum.

The incidence of neonatal deaths was numerically twice as high in the placebo group (2.7% vs. 6.0%), but the difference was not statistically significant. If miscarriages and stillbirths are added to the neonatal deaths, the overall fetal and neonatal mortality was similar in the 2 treatment groups (6.2% in the 17OHP-C group vs. 7.2% in the placebo group).

Table 8 Miscarriages, Stillbirths, and Neonatal Deaths

Pregnancy Outcome	17OHP-C N=306 n (%)	Placebo N=153 n (%)	Nominal P-value[A]
Miscarriages <20 weeks gestation	5 (1.6)	0	0.1746
Stillbirth	6 (2.0)	2 (1.3)	0.7245
Antepartum stillbirth	5 (1.6)	1 (0.6)	---
Intrapartum stillbirth	1 (0.3)	1 (0.6)	---
Neonatal deaths	8 (2.6)	9 (5.9)	0.1159
Total Deaths	**19 (6.2)**	**11 (7.2)**	**0.6887**

[A] No adjustment for multiple comparisons.

Source: Table 11-6 and Table 11-9, Final Report for Study 17-CT-002.

Division's Comments

- *The trend towards a benefit in the reduction of neonatal death is off-set by a trend toward an increase in the rates of miscarriage and possibly stillbirth associated with use of 17OHP-C, resulting in no net benefit regarding survival.*
- *Based on the data provided in Study 17P-CT-002, there is no indication that treatment with 17OHP-C will reduce overall fetal/neonatal mortality.*

4.5.4 Neonatal Outcomes

4.5.4.1 Neonatal Characteristics

Four hundred forty-six (446) live infants were delivered by 459 subjects with known delivery dates: 295 infants in the 17OHP-C group and 151 infants in the placebo group (Table 9).

Birthweight

The percentage of infants weighing <2500 g was significantly lower in the 17OHP-C group than in the placebo group (27.2% vs. 41.1%). The percentage of infants weighing <1500 g also was numerically (but not statistically) lower in the 17OHP-C group (8.6% vs. 13.9%). There were no differences between treatment groups in mean birthweight.

Apgar Scores
There were no differences between treatment groups in mean 1-minute and 5-minute Apgar scores.

Congenital Malformations
Nine (2.0%) infants overall had a major congenital malformation; the incidence rate was not different between treatment groups: 6 (2.0%) in the 17OHP-C group and 3 (2.0%) in the placebo group.

Admission to and Days in NICU
A smaller percentage of liveborn infants in the 17OHP-C group were admitted to the NICU compared with liveborn infants in the placebo group (27.8% vs. 36.4%). For live births, stay in the NICU ranged widely, from 0.1 - 194.8 days. The median stay in the NICU was numerically (but not statistically) shorter for the 17OHP-C group (9.1 vs. 14.1 days).

Hospital days were available for 285 infants in the 17OHP-C group and 140 infants in the placebo group. The difference in mean hospital days between treatment groups was not significant (8.7 vs. 13.3 days).

Table 9 Neonatal Outcomes in Study 17P-CT-002

Neonatal Outcome	17OHP-C	Placebo	Nominal P-value[A]
Number of subjects	310	153	--
Number of live births	295	151	--
Birthweight (g)			
Mean (SD)	2760 (859)	2582 (942)	0.0736
Min, Max	208, 4900	300, 4855	--
Birthweight <2500 g, n (%)	82 (27.2)	62 (41.1)	**0.0029**
Birthweight <1500 g, n (%)	26 (8.6)	21 (13.9)	0.0834
Head circumference			
Mean (SD)	32.5 (3.1)	32.0 (3.3)	0.0963
Min, Max	15.4, 37.5	21.5, 38.0	--
1 Minute Apgar			
Mean (SD)	7.5 (2.3)	7.3 (2.3)	0.2135
Min, Max	0, 9.0	0, 9.0	--
5 Minute Apgar			
Mean (SD)	8.3 (1.9)	8.3 (1.7)	0.1058
Min, Max	0, 10.0	0, 9.0	--
Major congenital malformation, n (%)	6 (2.0)	3 (2.0)	1.0000
Admitted to NICU or miscarriage/stillbirth/neonatal death, n (%)	93 (30.4)	57 (37.3)	0.1395
Admitted to NICU (live births), n (%)	82 (27.8)	55 (36.4)	**0.0434**
Days in NICU [B]			
Median	9.1	14.1	0.1283
Min, Max	0.1, 194.8	0.1, 147.0	--
Infant hospital days [C]			
Mean (SD)	8.7 (16.0)	13.3 (26.5)	0.3612
Min, Max	2, 123	2, 148	--

Birthweight and head circumference data were missing for some infants.

A: No adjustment for multiple comparisons

B: For neonatal deaths, days in the NICU were calculated until date of death. Days in NICU could not be determined for 3 patients in the 17OHP-C group and 2 patients in the placebo group.

C: Determined only for infants discharged alive.

Source: Table 11-7 Final Report for Study 17-CT-002.

4.5.4.2 Neonatal Morbidity other than Death for Live Births

The incidences of use of supplemental oxygen (15.4% vs. 24.2%), any type of intraventricular hemorrhage (IVH) (1.4% vs. 5.3%), and NEC (0% vs. 2.7%) were significantly lower in the 17OHP-C group than the placebo group (see Table 10). However, the incidence of severe IVH (Grades 3 or 4) was numerically higher in the 17OHP-C group (0.7% vs. 0.0%)

The incidences of the following neonatal morbidities, while not statistically different between treatment groups, were lower in the 17OHP-C group: BPD (1.4% vs. 3.3%); patent ductus

arteriosus (PDA) (2.4% vs. 5.4%); other intracranial hemorrhages (0.3% vs. 1.3%); and confirmed pneumonia (1.0% vs. 2.7%).

Composite neonatal morbidity was based on the number of neonates who died or experienced RDS, BPD, grade 3 or 4 IVH, proven sepsis, or NEC. The proportion of subjects who experienced the composite morbidity endpoint was numerically lower in the 17OHP-C group (11.9% vs. 17.2%), but the difference was not statistically significant.

Table 10 Neonatal Morbidity for Live Births

Morbidity	17P N=295 n (%)	Placebo N=151 n (%)	Nominal P-value [A]
Transient tachypnea	11 (3.7)	11 (7.3)	0.0990
Respiratory distress syndrome (RDS)	29 (9.9)	23 (15.3)	0.0900
Bronchopulmonary dysplasia (BPD)	4 (1.4)	5 (3.3)	0.1730
Persistent pulmonary hypertension	2 (0.7)	1 (0.7)	1.0000
Ventilator support	26 (8.9)	22 (14.8)	0.0616
Supplemental oxygen	***45 (15.4)***	***36 (24.2)***	***0.0248***
Patent ductus arteriosus	7 (2.4)	8 (5.4)	0.1004
Seizures	3 (1.0)	0	0.5541
Any intraventricular hemorrhage (IVH)	***4 (1.4)***	***8 (5.3)***	***0.0258***
Grade 3 or 4 IVH	2 (0.7)	0	0.5511
Other intracranial hemorrhage	1 (0.3)	2 (1.3)	0.2628
Retinopathy of prematurity	5 (1.7)	5 (3.3)	0.3164
Proven newborn sepsis	9 (3.1)	4 (2.6)	1.0000
Confirmed pneumonia	3 (1.0)	4 (2.7)	0.2330
Necrotizing enterocolitis (NEC)	***0***	***4 (2.7)***	***0.0127***
Composite Neonatal Morbidity Score [B]	*35 (11.9)*	*26 (17.2)*	*0.1194*

A: P-values have not been adjusted for multiple comparisons.

B: The composite neonatal morbidity measure counted any liveborn infant who experienced death, RDS, BPD, grade 3 or 4 IVH, proven sepsis, or NEC.

Source: Table 11-8, Final Report for Study 17P-CT-002.

Division's Comments

- *The Applicant did not adjust for multiple comparisons. Had such a correction been performed, it is unlikely that any of the listed morbidities would have been statistically lower in the 17OHP-C treatment group in this clinical trial.*
- *The composite neonatal morbidity score included neonatal death and the major morbid conditions of the neonate. Although the composite neonatal morbidity score was numerically lower in the 17OHP-C treatment group (11.9% vs. 17.2%), the difference did not reach statistical significance.*

4.5.5 Summary of Division's Assessment of the Efficacy of 17OHP-C in Study 17P-CT-002

The results from this study of 463 pregnant women with a history of prior spontaneous preterm deliveries show the following:

- The frequency of preterm birth $<37^0$ weeks gestation was significantly decreased in the 17OHP-C treatment group compared to that in the placebo group (37.1% vs. 54.9%). The reduction in preterm birth < 37 weeks was independent of race, number of qualifying preterm births, and gestational age of the qualifying preterm birth.
- The prolongation of pregnancy, defined as the time from randomization to delivery or date last pregnant, was significantly longer, by a mean of 6 days, in the 17OHP-C group compared to the placebo group. The mean gestational age at delivery was one week greater in the 17OHP-C group compared to the placebo group (36.2 vs. 35.2 weeks).
- Use of tocolytic therapy and cerclage placement were not significantly different between the 17OHP-C and placebo groups.
- The percentage of infants weighing <2500 g was lower in the 17OHP-C group compared with the placebo group (27.2% vs. 41.1%). The percentage of infants weighing <1500 g was not statistically different between the treatment groups.
- A smaller percentage of live births in the 17OHP-C group were admitted to the NICU (27.8% vs. 36.4%).
- Neonatal mortality was numerically lower in the 17OHP-C group, but the between-group difference was not statistically significant (2.6% vs. 5.9%).
- Five miscarriages (1.6% of pregnancies) occurred in the 17OHP-C group compared to no miscarriages (0%) in the placebo group.
- The rate of stillbirths was slightly higher in the 17OHP-C (2.0% vs. 1.3%), but the difference was not statistically significant.
- Composite neonatal morbidity (neonates with death, RDS, BPD, grade 3 or 4 IVH, proven sepsis, or NEC) was lower in the 17OHP-C group, but the between-group difference was not statistically significant (11.9 vs. 17.2).

4.6 Safety Outcomes

4.6.1 Collection of Safety Data

Studies 17P-IF-001 and 17P-CT-002 were conducted under an IND, but adverse events (AEs) were not captured in the typical manner used for studies designed to support a drug registration. Assessment of severity or relationship of AEs to study drug was not made for non-serious AEs. Adverse events that were considered serious and unexpected by the investigator were reported using the MFMU Network AE Form, which requested assessments of severity and relationship to study drug.

4.6.2 Deaths

4.6.2.1 *Maternal*

There were no maternal deaths in the trial.

4.6.2.2 Miscarriages, stillbirths, and neonatal deaths

There was a higher incidence of miscarriage and stillbirth in the 17OHP-C group (3.5% vs. 1.3%), but a lower incidence of neonatal deaths (2.6% vs. 5.9%). Neither of the between-group differences was statistically significant.

Miscarriages

Five (1.6%) subjects randomized to 17OHP-C had miscarriages, compared with no subjects randomized to placebo. Another 17OHP-C subject (Patient 004-035) had a spontaneous vaginal delivery of a nonviable fetus at 20^{1} weeks gestation, which was classified as a neonatal death; the infant had 1- and 5-minute Apgar scores of 1 and died the day of delivery due to extreme prematurity.

Two of the five subjects who had miscarriages had a clinical diagnosis of chorioamnionitis at the time of the miscarriage. Patient 008-114 miscarried after her 3rd injection of 17P, at 19^{1} weeks gestation. Patient 015-023 had a previous stillbirth, a previous miscarriage, and had a miscarriage on the day of her 2nd 17OHP-C injection, at 19^{1} weeks gestation.

Patient 015-014 had a previous stillbirth and during this pregnancy had bacterial vaginosis prior to randomization. She received 3 injections of 17OHP-C before experiencing pPROM at 18^{6} weeks gestation. She chose to terminate the pregnancy due to a poor prognosis for the infant. Although classified as an induced abortion on the AE form, the event was entered in the database as a miscarriage.

One subject, Patient 008-110, smoked a pack a day of cigarettes and used cocaine during the study pregnancy. After receiving a single injection of 17P, she experienced a miscarriage at 18^{2} weeks gestation.

Only one of the five subjects who had a miscarriage had no identifiable factor that might have contributed to the miscarriage. However, prior to entering the study, this subject, Patient 004-048, had an emergency room visit for a threatened abortion at 9^{4} weeks gestation. She was randomized to 17OHP-C at 17^{3} weeks gestation and received her only injection of 17OHP-C on that day. Five days later, she experienced pPROM and had a spontaneous vaginal delivery of a nonviable infant.

Division's Comment

- *Although the Applicant notes that infection appears more likely to be contributory to miscarriage than does exposure to 17OHP-C, the rate of chorioamnionitis and vaginitis in placebo women (none of whom miscarried) was not significantly lower.*

Stillbirths

There were a total of 8 stillbirths, 6 occurring in the 17OHP-C group and 2 in the placebo group. The difference in incidence of stillbirths was not statistically significant (2.0% for 17OHP-C vs. 1.3% for placebo).

Two of the stillbirths, one in each treatment group, occurred intrapartum. Neither subject had a prior stillbirth. Subject 023-007 started 17OHP-C at 18^{5} weeks gestation of her 4th pregnancy and received 3 injections with no AEs. She had nothing in her obstetrical history that could explain the stillbirth at 21^{0} weeks gestation. Subject 008-060 started placebo at 18^{4} weeks gestation. She had bacterial vaginosis prior to randomization. She received 5 injections of placebo with no AEs, and then developed preeclampsia at 23^{6} weeks gestation

with symptoms consistent with placental abruption and labor was induced; a stillborn fetus was delivered.

Six of the stillbirths occurred as fetal deaths in utero (5 in the 17OHP-C arm; one in the placebo arm). Three 17OHP-C subjects (008-102, 015-022, and 017-011) had bacterial vaginosis or Trichomonas vaginalis during the study pregnancy prior to randomization. Subject 014-012 in the 17OHP-C group had a clinical diagnosis of chorioamnionitis during the pregnancy. These infections may have played some role in causing the stillbirths. Subject 018-024 in the 17OHP-C group had no identifiable factor in her obstetrical history or study data that could have contributed to the stillbirth. The placebo subject, Subject 013-005, had a urinary tract infection before randomization and was a smoker (10 cigarettes/day).

Division's Comment

- *Data on second trimester miscarriage rates also are available from 4 studies reported in a meta-analysis of published studies.[28] Data in the meta-analysis publication showed a trend toward an increased risk of miscarriage in the 17OHP-C arms as compared to placebo (odds ratio of 1.30, with 95% confidence interval of 0.61 – 2.74).*
- *The results of the current clinical trial, along with the meta-analysis, demonstrated a trend toward increased second trimester miscarriage.*

Neonatal Deaths

The incidence of neonatal death was twice as high in the placebo group, with 9 deaths (5.9% of births) occurring in the placebo group, as compared to 8 in the 17OHP-C group (2.6%). This did not reach statistical significance. The gestational ages at delivery of these infants ranged from 20.3 to 28.1 weeks in the placebo group and from 20.1 to 35.1 weeks in the 17OHP-C group. The neonatal death in the 35-week delivery in the 17OHP-C group occurred in an infant delivered by emergency caesarian section following uterine rupture. Excluding this infant, the gestational age at the time of the delivery of the neonatal deaths was similar between groups.

Division's Comment

- *The similar gestational ages at delivery of the neonatal deaths in the 2 groups suggests that the gestational age-adjusted neonatal death rate would be similar for each group. This further suggests that the decreased neonatal death rate in the 17OHP-C group is attributable to a lower proportion of early preterm deliveries, rather than a difference in the condition of the delivered neonates.*

4.6.3 Congenital Anomalies

The incidence of congenital malformations was 2% in both treatment groups. The 6 cases in the 17OHP-C group included 2 congenital genitourinary anomalies (a male with obstructive defects of the renal pelvis and ureter and a female with a hydrocele of the tunica vaginalis), 2 infants with congenital cardiovascular anomalies (cardiomegaly/left ventricular diverticulum/ pericardial defect and one reported as "other anomalies of the circulatory system"), one infant with polydactyly and talipes calcaneovarus and one with congenital flat feet. The 3 cases in the placebo group were an infant with a congenital cardiovascular anomaly (stenosis and other anomalies of the circulatory system) and polydactyly, one with a

congenital genitourinary anomaly (anomalies of the bladder and urethra), and one with talipes equinovarus.

Division's Comment

- *The number and type of congenital anomalies appear evenly distributed over the treatment arms. This rate of anomalies is consistent with the background rate for congenital anomalies in the general population of 2-3%.*

4.6.4 Non-Fatal Serious Adverse Event Reports

Four subjects, all of whom received 17OHP-C, had non-fatal AEs that triggered the submission of a serious unexpected adverse event report.

Patient 002-026 had a pulmonary embolus after delivery. The subject was randomized to 17OHP-C at 19^{4} weeks gestation and received 17 injections of 17OHP-C before delivery. She had a labor visit between the 8^{th} and 9^{th} injections and again between the 15^{th} and 16^{th} injections of study drug. She experienced significant antepartum bleeding during the second labor visit and had a positive lupus anti-coagulant test, but continued in the study. She had no symptoms of thromboembolic events during the pregnancy. Eight days after delivery at 36^{4} weeks, the subject experienced a pulmonary embolus, which was successfully treated and did not result in any sequelae.

Patient 013-021 reported a knot at the injection site on her right hip, which was very sore, after the 8th injection of 17P. She was diagnosed with cellulitis and started on penicillin. The subject requested to remain in the study and had a spontaneous PTD at 31^{4} weeks gestation.

Patient 017-016 delivered a male infant at 37^{5} weeks gestation with small penis and testes. An ultrasound of the scrotum revealed infarcted testicles secondary to intrauterine torsion. Human chorionic gonadotropin, congenital hypothyroidism, and follicle stimulating hormone, and chromosome testing were done and found to be normal. The infant was diagnosed as possibly having hypogonadism.

Patient 014-012 had a stillbirth at 21^{1} weeks gestation, and developed postpartum hemorrhage and respiratory distress after delivery. The subject was intubated and given multiple transfusions of red blood cells before being discharged to specialty care. The subject continued on antibiotics for endometritis and excessive surgical manipulation.

Division's Comment:

- *A causal association of these 4 maternal serious AEs with exposure to 17OHP-C is unlikely.*

4.6.5 Common Adverse Events

The most common AEs in both treatment groups were injection site reactions, reported by 42.3% of 17OHP-C subjects and 38.6% of placebo subjects. The types of injection site reactions did not differ between the treatment groups, except for injection site swelling, which occurred with a significantly greater incidence in the 17OHP-C group compared with the placebo group (17.1% vs. 7.8%). Adverse events by preferred terms that occurred in >2% of subjects in either treatment group are shown in Table 11.

Table 11 Adverse Events that Occurred in >2% of Subjects in either Treatment Group

Preferred Term [A]	17P N=310 n (%)	Placebo N=153 n (%)
Injection site pain	108 (34.8)	50 (32.7)
Injection site swelling [B]	53 (17.1)	12 (7.8)
Urticaria	38 (12.3)	17 (11.1)
Pruritus	24 (7.7)	9 (5.9)
Injection site pruritus	18 (5.8)	5 (3.3)
Nausea	18 (5.8)	7 (4.6)
Contusion	17 (5.5)	14 (9.2)
Injection site nodule	14 (4.5)	3 (2.0)
Vomiting	10 (3.2)	5 (3.3)
Death [C, D]	8 (2.6)	9 (5.9)
Anorexia	5 (1.6)	6 (3.9)
Injection site irritation	4 (1.3)	5 (3.3)
Abdominal pain	3 (1.0)	4 (2.6)

[A] Patients reporting a particular AE more than once were counted only once for that AE. AEs were coded using MedDRA Version 8.0.

[B] Incidence in 17OHP-C group was significantly higher (p>0.05) than placebo group, based on a chi-square test.

[C] Death included only neonatal deaths.

[D] For safety assessments, the incidence of neonatal death was based on all randomized patients, so the percentages are slightly lower than those reported for the efficacy assessment based on liveborn infants.

Source: Table 12-2, Final Report for Study 17-CT-002.

Infections were not recorded as AEs during the study, but were captured indirectly if they resulted in antibiotic use. The incidence of any vaginal/cervical infection was greater in the 17OHP-C group (21.6%) as compared to the placebo group (15%). Incidences in the 17OHP-C and placebo groups, respectively, of bacterial vaginosis (8.7% vs. 5.2%) and trichomonas (3.9% vs. 1.3%) did not differ significantly.

4.6.6 Adverse Events That Led to Discontinuation of Study Drug

The rate of early discontinuations from treatment with study drug due to AEs was comparable in the 2 treatment groups and the AEs leading to discontinuation were not notably different. Seven (2.2%) subjects in the 17OHP-C group and four (2.6%) subjects in the placebo group either discontinued or were withdrawn by the investigator from study drug due to AEs.

The principal AEs that led to discontinuation from treatment in the 17OHP-C and placebo groups are listed by subject in Table 12:

Table 12 Adverse Events Leading to Treatment Discontinuation (Study 17P-CT-002)

Patient ID	Treatment Group	Adverse Event	Gestational Age at Discontinuation
002-024	17OHP-C	Urticaria	23.3 weeks
004-018	17OHP-C	Soreness at injection site	23.3 weeks
008-055	placebo	Pruritus (head to toe)	20.1 weeks
011-027	17OHP-C	Arthralgia/Severe Joint Pain	19.6 weeks
015-033	placebo	Swelling at injection site/Pruritus	30.6 weeks
018-018	placebo	Urticaria	26.1 weeks
019-015	17OHP-C	Urticaria	31.1 weeks
020-026	17OHP-C	Weight Gain	26.3 weeks
020-044	17OHP-C	Urticaria	24.3 weeks
020-060	17OHP-C	Red welt at injection site	20.5 weeks
025-001	placebo	Pruritus	34.3 weeks

Source: Section 16.2, Listing 7.5, Final Report for Study 17P-CT-002

Another subject in the 17OHP-C group was listed as being withdrawn early by the investigator due to pPROM, which was not considered an AE in this study. Four subjects in the 17OHP-C group were lost to follow up, and one of these 4 subjects reported swelling at the injection site at the last 2 visits before being lost to follow up. The other 3 subjects who were lost to follow up had no AEs reported.

A placebo subject was also withdrawn early by the investigator due to pPROM.

Twenty-eight other subjects discontinued study drug early due to non-clinical reasons: 19 in the 17OHP-C group and 9 in the placebo group. No other information was provided on the CRF as to why the subject discontinued. Of the 19 subjects in the 17OHP-C group, 12 had no recorded AEs. Of the remaining 7 subjects, 4 had AEs within 2 visits of discontinuation, and therefore, without additional information as to the reason for discontinuation, the role of an AE in the decision to discontinue can not be excluded. The AEs reported by these subjects prior to discontinuation were injection site reactions (n=3) and diarrhea, vomiting, and loss of appetite (n=1 for each).

Of the 9 subjects in the placebo group who discontinued for non-clinical reasons, 6 had no recorded AEs. Of the remaining 3 subjects, one subject reported itching (pruritus) at the time of discontinuation.

Division's Comment:

- *The Applicant computed a worst-case scenario by adding the five 17OHP-C subjects and the one placebo subject who experienced AEs shortly before discontinuation/loss to follow-up to the group of subjects who discontinued due to AEs. By this conservative estimate of the incidence of discontinuation due to AEs, the incidence is still similar between the treatment groups (3.9% vs. 3.3%).*
- *The majority of AEs that clearly or possibly led to early discontinuation were injection site reactions, which occurred with both 17OHP-C and placebo. Two subjects, one in each treatment group, had possible allergic reactions, which have been reported previously for 17OHP-C.*

4.6.7 Pregnancy Complications and Maternal Outcomes

The incidence of maternal pregnancy complications (gestational diabetes, oligohydramnios, significant antepartum bleeding, preeclampsia/gestational hypertension, abruption, confirmed clinical diagnosis of chorioamnionitis, or cerclage placement) did not differ between the treatment groups (Table 13). The most common pregnancy complications (>5% of subjects in either treatment group) were preeclampsia or gestational hypertension (8.8% vs. 4.6%) and gestational diabetes (5.6% vs. 4.6%).

Overall, 70 subjects were admitted for preterm labor (PTL), other than the delivery admission, with similar rates in the 2 treatment groups: 16.0% of 17OHP-C subjects and 13.8% of placebo subjects. The mean length of hospital stay for the mothers was not different between the treatment groups (3.1 vs. 3.7 days).

Table 13 Pregnancy Complications

Complication or Outcome	17P N=306 n (%)	Placebo N=152 n (%)
Hospital or labor/delivery admission for PTL (other than the delivery admission)	49 (16.0)	21 (13.8)
Gestational diabetes	17 (5.6)	7 (4.6)
Oligohydramnios	11 (3.6)	2 (1.3)
Significant antepartum bleeding	6 (2.0)	3 (2.0)
Preeclampsia or gestational hypertension	27 (8.8)	7 (4.6)
Abruption	5 (1.6)	4 (2.6)
Confirmed clinical chorioamnionitis	11 (3.6)	5 (3.3)
Cerclage placement	5 (1.6)	2 (1.3)
Other complication	8 (2.7)	5 (3.3)

Source: Table 12-3 Final Report for Study 17-CT-002.

4.6.8 Laboratory Findings

No blood samples for routine laboratory tests were collected.

4.6.9 Summary of Overall Safety

This study exposed 310 pregnant women to 17OHP-C, with an average of 14.1 injections, compared with 153 pregnant women who received an average of 13.7 injections of placebo. Comparing the safety profile in each group:

- No maternal deaths occurred in either treatment arm.
- The frequency of both miscarriage and stillbirth was higher in the 17OHP-C group, although not statistically significantly different.
- The incidence of neonatal death, also not statistically significantly different between the 2 treatment arms, occurred at more than twice the rate in the placebo group.
- The incidence of congenital malformations was consistent with the normal background rate of 2% in both treatment groups, and the types of anomalies were similar.

- Twenty-nine (9.4%) subjects or their infants in the 17OHP-C group and 15 (9.8%) subjects or their infants in the placebo group experienced at least one serious or unexpected AE. The most common serious AE was fetal or neonatal death (miscarriages, stillbirths, and neonatal deaths). Maternal serious AEs occurred in four 17OHP-C subjects, but were not clearly related to study drug exposure.
- The overall incidence of AEs, including the most common AE (injection site reaction) was similar in the 17OHP-C and the placebo groups. The incidence of injection site swelling was significantly higher in the 17OHP-C group than the placebo group. All other injection site reactions occurred at comparable rates in the treatment groups.
- Early discontinuations due to AEs occurred at a comparable rate in the 17OHP-C and placebo groups, and were most often associated with injection site reactions.
- The incidence of pregnancy complications and the mean length of hospital stay for mothers did not differ between the 2 treatment groups.

5 SUPPORTIVE CLINICAL TRIAL

Study 17P-IF-001: "A Randomized Trial of 17α-Hydroxyprogesterone Caproate (Initial Formulation) for Prevention of Preterm Birth in High Risk Women"

5.1 Background

This study was designed and initiated in 1998 by the NICHD MFMU Network to evaluate the use of 17OHP-C for the prevention of recurrent preterm births. In February 1999, the manufacturer of study drug issued a recall, at the request of the Food and Drug Administration (FDA), because of violations of manufacturing practices that may have jeopardized the validity and potency of the product. On March 15, 1999, the study was terminated. At the time of study termination, only 150 of the proposed 500 subjects had been enrolled into the study. Eighty-six subjects completed the treatment regimen before the study was stopped (57 [61%] of the 17OHP-C subjects and 29 [52%] of the placebo subjects). The study was freshly started at a later date as Study 17P-CT-002 (see Section 4) when a new manufacturer was identified.

5.2 Overall Study Design

The study design for Study 17P-IF-001 was identical to that of Study 17P-CT-002 and is described in detail in Section 4.3. Briefly, the study was a randomized, placebo-controlled, efficacy and safety study of 17OHP-C in pregnant women, beginning at 16^0 to 20^6 weeks gestation, who had a history of spontaneous preterm birth, defined as delivery from 20^0 to 36^6 weeks gestation following spontaneous PTL or pPROM.

Qualifying subjects were randomized in a 2:1 ratio to 17OHP-C or placebo (castor oil with 46% benzyl benzoate and 2% benzyl alcohol). Study drug was administered weekly by intramuscular injection through 36^6 weeks gestation or delivery, whichever occurred first. The primary efficacy outcome was birth prior to 37^0 weeks (as determined by the gestational age established during the study).

5.3 Findings

5.3.1 Subject disposition

A total of 150 subjects were randomized, 94 to 17OHP-C and 56 to placebo. Sixty-five subjects randomized to 17OHP-C and 39 subjects randomized to placebo either completed treatment with study drugs or were withdrawn prematurely because of reasons other than recall of study drugs. Fifty-seven (61%) of subjects in the 17OHP-C group and 29 (52%) in the placebo group completed treatment through 36^6 weeks or delivery.

Among the subjects not impacted by recall of study drug, the reasons for not completing treatment in the 17OHP-C group (other than for recall of study drug) were adverse event (n = 1), withdrawal for non-clinical reasons (n = 6), and lost to follow up (n = 1). The reasons for not completing treatment in the placebo group (other than for recall of study drug) were adverse event (n = 2), withdrawal for non-clinical reasons (n = 6), and lost to follow up (n = 2).

5.3.2 Efficacy Findings

5.3.2.1 Primary Efficacy Endpoint

The incidence of delivery $<37^0$ weeks gestation for the ITT population, the population for which data were available (all subjects other than those lost to follow up) and those subjects whose participation was not prematurely terminated because of recall of study drug are listed in Table 14. For each analysis population, the percentage of subjects with a delivery of $<37^0$ weeks gestation was numerically higher in the 17OHP-C treatment group. None of the differences were statistically different.

Table 14 Percentage of Subjects with Delivery $<37^0$ Weeks Gestation

Analysis Population	17P		Placebo	
	N	n (%)	N	n (%)
ITT population	94	39 (41.5)	56	20 (35.7)
All available data	93	38 (40.9)	54	18 (33.3)
Not withdrawn due to study termination	**65**	**28 (43.1)**	**39**	**15 (38.5)**

ITT population was all randomized subjects. Subjects with missing outcome data were classified as having a preterm birth $<37^0$ weeks (treatment failure).

Source: Table 9-3, pg 21, abbreviated Final Report for Study 17P-IF-001.

Division's Comment

- *The data obtained from the analysis population identified as "not withdrawn due to study termination" is of most value since all subjects in this population had the opportunity to complete a full course of treatment. However, because the potency and overall quality of the study drugs could not be assured, the efficacy data obtained from this prematurely terminated clinical trial is of limited value and must be interpreted with caution. The findings from this trial do not suggest any benefit of 17OHP-C (at the uncertain dose that was administered) in reducing the percentage of subjects with a delivery $<37^0$ weeks gestation.*

- *In the "not withdrawn due to study termination" analysis population, the percentage of subjects with a delivery <37^0 weeks gestation was 38.5% in the placebo group. This rate of premature birth is close to that which the NIH used in their sample size calculations for both this study and study 17P-CT-002. This rate also is close to rates for high risk subjects reported in the literature. The percentage of subjects with a delivery <37^0 weeks gestation in the placebo group of Study 17P-CT-002, however, was considerably higher ─ 54.9%. The difference in the rates of premature birth in the placebo arms of the 2 clinical trials (38.5% vs. 54.9%) is surprising since both studies were conducted at the same clinical sites in close temporal proximity.*

5.3.2.2 Secondary Efficacy Outcomes

Miscarriages, Stillbirths, and Neonatal Deaths

The number and percentages of miscarriages, stillbirths, and neonatal deaths in the ITT population are listed by treatment group in Table 15.

Table 15 Number of Miscarriages, Stillbirths, and Neonatal Deaths

Fetal/Neonatal Deaths	17P N=93	Placebo N=54
Miscarriages	1 (1.1)	1 (1.9)
Stillbirths	1 (1.1)	2 (3.7)
Neonatal deaths	2 (2.2)	0
Total	**4 (4.4)**	**3 (5.9)**

Source: Table 9-8, pg 28, abbreviated Final Report for Study 17P-IF-001.

Division's Comment

- *Although this study did not demonstrate any overall benefit for treatment with 17OHP-C in terms of reduction in overall mortality, there was no trend toward an increased rate of miscarriages in the 17OHP-C group as was seen in Study 17P-CT-002.*

5.3.3 Safety Findings

5.3.3.1 Deaths and Discontinuations because of Adverse Events

There were no maternal deaths. Four subjects, 2 in the 17OHP-C group and 2 in the placebo group, discontinued study drug early due to AEs. One 17OHP-C subject discontinued after the first injection due to an injection site rash, and the other 17OHP-C subject discontinued after the ninth 17OHP-C injection due to urticaria, swelling, and redness at the injection site. One placebo subject discontinued after the first injection due to vomiting, urticaria, and facial swelling and the other placebo subject discontinued after the seventh injection due to urticaria.

5.3.3.2 Common Adverse Events

Of the 150 subjects, 92 (61.3%) reported 368 AEs during the study; 60 (63.8%) subjects in the 17OHP-C group reported 237 AEs, and 32 (57.1%) subjects in the placebo group reported 131 AEs. The most commonly reported AEs were injection site reactions, which occurred at a comparable rate in the 2 treatment groups (52.1% in 17OHP-C vs. 46.4% in the

placebo group). Adverse events that occurred in >2% of subjects in either treatment group are shown in Table 16 by preferred terms in descending order of incidence in the 17OHP-C group.

Table 16 Adverse Events that Occurred in >2% of Subjects in a Treatment Group

Preferred Term	17OHP-C N=94 n (%)	Placebo N=56 n (%)
Injection site pain	41 (43.6)	24 (42.9)
Injection site swelling	15 (16.0)	6 (10.7)
Urticaria	13 (13.8)	7 (12.5)
Contusion	9 (9.6)	6 (10.7)
Injection site pruritus	7 (7.4)	5 (8.9)
Pruritus	4 (4.3)	2 (3.6)
Injection site nodule	3 (3.2)	4 (7.1)
Abdominal pain	3 (3.2)	2 (3.6)
Nausea	2 (2.1)	1 (1.8)
Injections site erythema	2 (2.1)	0
Edema	2 (2.1)	0
Diarrhea	2 (2.1)	0
Death (neonatal)	2 (2.1)	0
Stillbirth	1 (1.1)	2 (3.6)
Dizziness	0	2 (3.6)

Source: Table 19-2, pg 34, abbreviated Final Report for Study 17P-IF-001.

5.3.3.3 Pregnancy Complications

The most common pregnancy complications in the 17OHP-C group (other than admission for preterm labor not related to the delivery admission) were gestational diabetes (8.6% 17OHP-C vs. 0% placebo) and preeclampsia or gestational hypertension (6.5% 17OHP-C vs. 3.8% placebo) (see Table 17). There was almost double the rate of hospitalization for preterm labor (other than the delivery admission) in the placebo group as compared to the 17OHP-C group.

Table 17 Pregnancy Complications

Complication	17OHP-C N=93 n (%)	Placebo N=53 n (%)
Hospital or labor/delivery admission for preterm labor (other than the delivery admission)	10 (10.8)	11 (20.8)
Gestational diabetes	8 (8.6)	0
Oligohydramnios	2 (2.2)	1 (1.9)
Significant antepartum bleeding	4 4.3)	4 (7.5)
Preeclampsia or gestational hypertension	6 (6.5)	2 (3.8)
Abruption	2 (2.2)	2 (3.8)
Confirmed clinical chorioamnionitis	2 (2.2)	0
Cerclage placement	0	1 (1.9)
Other complication	2 (2.2)	1 (2.0)

Source: Table 10-1, pg 38, abbreviated Final Report for Study 17P-IF-001.

Division's Comment

Comparing the safety profile in each group:

- No maternal deaths occurred in either treatment arm.
- The frequency of miscarriage, stillbirth, and neonatal death was not different in the 2 arms.
- The overall incidence of AEs, including the most common AE (injection site reaction) was similar in the 17OHP-C and the placebo groups. The incidence of injection site swelling was numerically higher in the 17OHP-C group than the placebo group. All other injection site reactions occurred at comparable rates in the treatment groups.
- Early discontinuations due to AEs occurred at a comparable rate in the 17OHP-C and placebo groups, and were most often associated with injection site reactions.
- The incidence of pregnancy complications was not different between the 2 treatment groups.
- The rate of admission for preterm labor, aside from the delivery hospitalization, was greater in the placebo group.

6 STUDY 17P-FU (FOLLOW-UP SAFETY STUDY)

6.1 Description of the Protocol

Infants born to women enrolled in Study 17P-CT-002, and who survived to be discharged from the nursery, were eligible for participation in the follow-up study, known as Study 17P-FU.

Instruments and Procedures

Assessment of the children's longer-term outcomes was performed using the following instruments and procedures:

- The primary endpoint was determined based upon the Ages and Stages Questionnaire (ASQ), completed by the parent or guardian
- Secondary endpoints were based upon items evaluated through use of
    - A Survey Questionnaire, administered by study personnel to the parent
    - Physical examination by a study pediatrician

The ASQ is composed of 19 questionnaires, each corresponding to a specific age window between 4 months and 5 years, and each containing 30 developmental items addressing five areas: communication, gross motor, fine motor, problem solving, and personal-social. The instrument was developed on a population including both children considered to be at risk for developmental problems and a normative sample of full term children with no health or developmental concerns. It has been validated against common professional assessment scales, including the Bayley Scales of Infant Development and the McCarthy Scales of Children's Abilities. The questionnaires are designed to identify young children who are in need of further evaluation and early intervention services. Cutoff points, generally corresponding to scores falling 2 standard deviations (SD) below the mean for the combined "at risk" and normal population, were generated for each of the five developmental domains assessed.

The Survey Questionnaire used in this study was derived from questions that were developed and reportedly validated by the following sources: the 2001 Child Health Supplement of the National Health Interview Survey, the 1991 National Maternal and Infant Health Survey, Early Childhood Longitudinal Survey (Department of Education), and the Avon Longitudinal Study of Parents and Children. This questionnaire was not formatted for self-administration; therefore it was administered by study personnel during the clinic visit. The Survey Questionnaire included evaluation of:

- Overall activity level and motor control, compared to age mates of the child, as measured by questions from the Early Childhood Longitudinal Study, Kindergarten (ECLSK), answered by the parent. If a perceived problem was reported by the parent, further questioning determined whether a professional evaluation and diagnosis had been made.
- Vision or hearing problems, assessed by questions from the National Health Interview Survey (NHIS), answered by the parent.
- Assessment of height, weight and head circumference, compared to reference curves generated by the Centers for Disease Control (CDC).
- Gender-specific behavior, assessed by the Pre-School Activities Inventory (PSAI).
- Diagnosis by a healthcare provider of cerebral palsy, asthma, allergic disorders, sensory disorders and neurodevelopmental disorders including attention deficit hyperactivity disorder (ADHD).

Division's Comment

- *Although the ECLSK was developed for use with children from kindergarten to fifth grade, the motor control and activity questions were reviewed by an NICHD*

developmental psychologist, who reportedly determined that they were appropriate for children as young as 2. The basis for this conclusion was not provided.

A general physical examination was conducted by a pediatrician or nurse practitioner at the study center, and included measurements of the child's current weight, height, head circumference, and blood pressure, as well as the documentation of any major abnormality. In addition, a part of the examination was specifically directed toward the identification of genital abnormalities. If the child had a physical examination within the last year, and the parent/guardian was unable to bring the child in for a visit, the information from that previous physical examination was entered into the study database. In these cases, the medical record of the child was abstracted by an NICHD pediatrician.

Following IRB approval, MFMU Network study personnel attempted to locate the women who participated in Study 17P-CT-002. If the mother who was enrolled in Study 17P-CT-002 could not be found, but her child could be located, the child's father or guardian could enroll the child in this study.

The nurse used a standardized script to request consent to participate. If the parent was willing to allow the child to participate, the nurse obtained informed consent by mail. She also made arrangements for the child to visit the Network center accompanied by the parent. In addition, the ASQ was mailed to the parent with instructions to bring the completed form to the visit. If the parent was unable to attend a follow-up visit, the research nurse administered the Survey Questionnaire by telephone, and asked the parent to mail back the completed ASQ.

The following procedures were conducted at the study visit:

- Administration of the Survey Questionnaire
- Physical examination
- Completion of the ASQ, if not done prior to the study visit

Parents were instructed to complete the ASQ based on the age of the child at the follow-up visit. The ASQ recommends using gestational age-corrected age only until 24 months and since all children to be evaluated were at least 2 years of age, corrected age was not used in this study. The completed ASQ was scored by the Biostatistical Coordinating Center (BCC) and results were sent back to the study nurse. If a child fell below a pre-established cutoff (below 2 SD from the mean) in at least one of the five developmental domains on the ASQ, the study nurse was to inform the parent/guardian that the child might need additional evaluation in the particular developmental area.

At the time of enrollment in Study 17P-FU, some of the mothers had already been informed of their treatment assignment in Study 17P-CT-002. If they had not, the treatment group was not revealed before the follow-up assessments. Less than 10% of the mothers were informed of their treatment (8.3% in the 17OHP-C group and 7.1% in the placebo group). The physicians or nurse-practitioners who performed the physical examinations were blinded to the treatment group assignment of the mother.

6.2 Inclusion/Exclusion Criteria

Inclusion Criteria

1. Maternal enrollment in the Study 17P-CT-002 conducted at one of the 14 Network centers in the fourth MFMU Network cycle (2001-present). As the composition of the MFMU changes over time, only women initially enrolled at a site that remained in the Network were eligible for the follow-up study.
2. Infant discharged alive from birth hospitalization.

Exclusion Criteria

No exclusion criteria were defined in the protocol.

6.3 Primary and Secondary Endpoints

The primary objective of the study was to determine if there were differences in achievement of developmental milestones between children whose mothers received 17OHP-C and those who received placebo in Study 17P-CT-002, as measured by the ASQ. The primary endpoint was the proportion of children from each treatment arm who fell below a specified cut-off on at least one of the five developmental areas measured on the ASQ.

The secondary objectives of the study were to determine if differences existed between children whose mothers received 17OHP-C and those who received placebo in Study 17P-CT-002 in the following factors:

- Gender-specific play
- Physical growth (height and weight)
- Activity levels
- Motor control
- Vision or hearing difficulties
- Physician- or other health provider-diagnosed conditions, such as asthma, allergic disorders, sensory disorders, and neurodevelopmental disorders such as attention deficit hyperactivity disorder (ADHD), as reported on the Survey Questionnaire

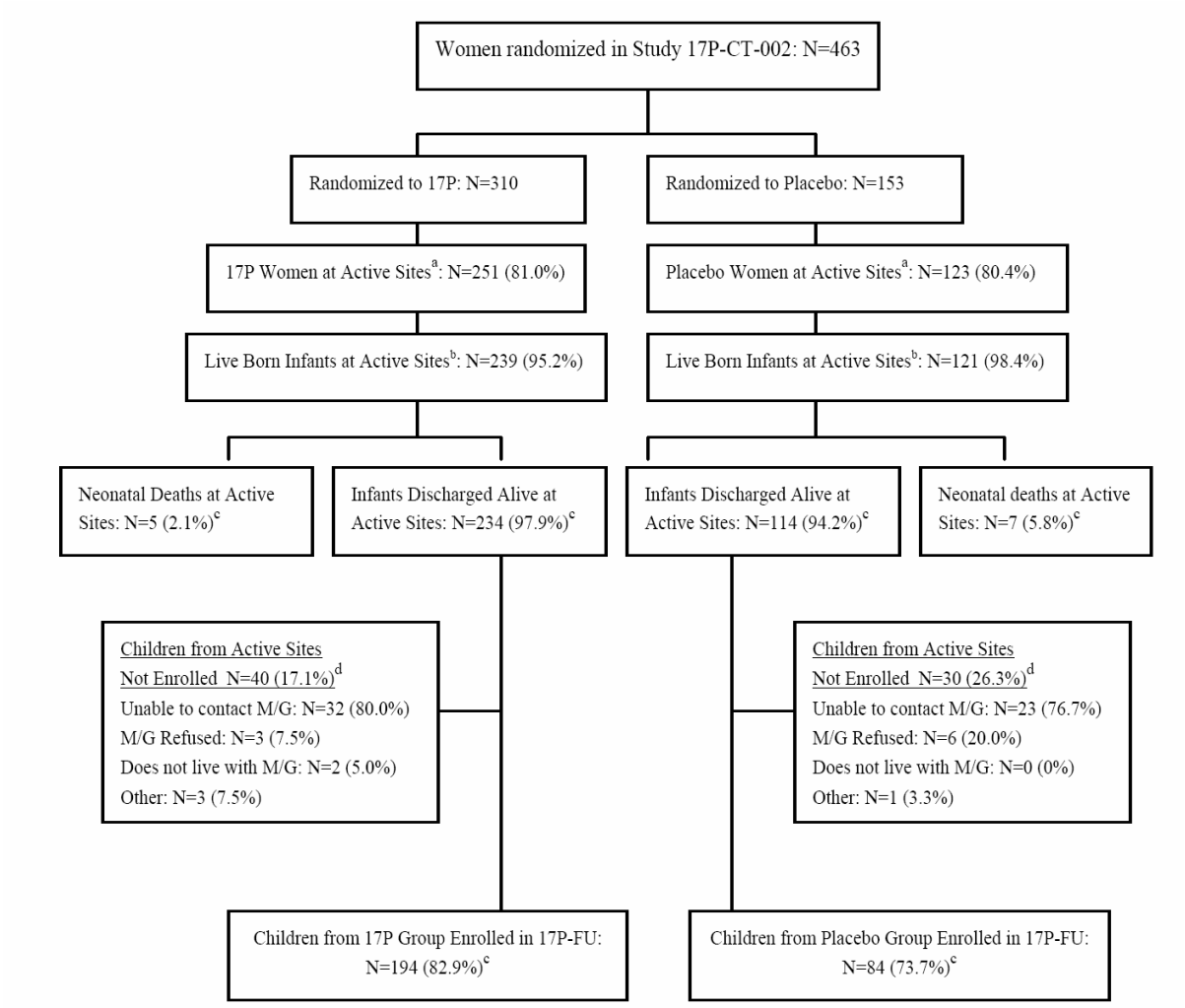
6.4 Subject Disposition

Figure 3 shows the disposition of infants born alive to mothers in Study 17P-CT-002. A total of 463 women were randomized to study drug; 310 women received 17OHP-C and 153 women received placebo. Of those women, a total of 374 women (251 [81.0%] of the 17OHP-C women and 123 [80.4%] of the placebo women) were enrolled at one of the 14 study sites still active in the MFMU Network at the start of this follow-up study. These women had a total of 360 live born infants, representing 74% of the 446 live births in Study 17P-CT-002. Twelve infants from the active sites died before discharge from the birth hospitalization, five (2.1%) of the 239 in the 17OHP-C group and seven (5.8%) of the 121 in the placebo group. There were no deaths following discharge from the nursery in children from the subset of mothers who were able to be located.

Of 348 eligible children, 278 (79.9%) were enrolled in Study 17P-FU. The percentage of eligible children who were enrolled in Study 17P-FU was greater in the 17OHP-C group (82.9% of the 17OHP-C-exposed vs. 73.7% of placebo-exposed). Inability to contact the

parent was the primary reason children were not enrolled. A greater proportion of placebo-treated mothers refused to allow their child to participate (5% of eligible placebo mothers vs. 1% of 17OHP-C-treated mothers).

Figure 3 Disposition of Subjects in Follow Up Study 17P-FU



Abbreviations: M/G = mother/guardian

[a] An active study site was a clinical center participating in the MFMU Network at the time Study 17P-FU was conducted.

[b] Percentages were based on the number of patients from active study sites.

[c] Percentages were based on the number of live born infants in Study 17P-CT-002 from active study sites.

[d] Percentages were based on the number of live born infants in Study 17P-CT-002 discharged from birth hospitalization from active study sites.

Source: Section 10.1, Figure 10-1, Final Report for Study 17P-FU.

6.5 Demographics and Other Baseline Characteristics

6.5.1 Demographics

The children ranged in age from 30 to 64 months at the time of enrollment. The mean age was similar for the 2 treatment groups (47.2 months for 17OHP-C vs. 48.0 months for the

placebo group), as was the distribution across the race/ethnic groups, which was assigned based on the mother's race or ethnicity. The majority of children were of African American descent (54.1% in the 17OHP-C group and 56.0% in the placebo), with children of Hispanic descent comprising 14.9% (17OHP-C) to 17.9% (placebo). Approximately one-fourth of the children were Caucasian. The 17OHP-C group had 58.3% male children compared with 47.6% in the placebo group.

6.5.2 Neonatal Outcomes of Enrolled Children

The neonatal outcomes of the enrolled children are listed in Table 18.

The gestational age at delivery ranged from 25.0 to 41.9 weeks, with a mean gestational age of 37.3 weeks in the 17OHP-C group and 36.2 weeks in the placebo group. This was slightly greater than the mean gestational ages observed in the total population in Study 17P-CT-002 (36.2 weeks for 17OHP-C vs. 35.2 for placebo).

Birthweight ranged from 714 - 4900 g in the 17OHP-C group and 615 - 4855 g in the placebo group. The 17OHP-C group had a lower percentage of infants with birthweight <2500 g (21.8% vs. 34.5%) and <1500 g (4.7% vs. 8.3%). The mean and range of APGAR scores were comparable between the 2 treatment groups.

Table 18 Neonatal Outcomes of Enrolled Children

Characteristic	17OHP-C	Placebo
Gestational age at delivery (wks)	N=194	N=84
Mean (SD)	37.3 (3.2)	36.2 (3.7)
Min, Max	25.0, 41.7	25.1, 41.9
Birthweight (g)	N=193	N=84
Mean (SD)	2,914 (707.8)	2,756.7 (813.7)
Min, Max	714, 4900	615, 4855
Birthweight <2500 g, n (%)	42 (21.8)	29 (34.5)
Birthweight <1500 g, n (%)	9 (4.7)	7 (8.3)
Head Circumference (cm)	N=188	N=82
Mean (SD)	32.8 (2.5)	32.2 (3.2)
Min, Max	23.0, 37.5	21.5, 38.0
1 Minute APGAR	N=191	N=84
Mean (SD)	7.8 (1.6)	7.6 (1.7)
Min, Max	1.0, 9.0	1.0, 9.0
APGAR <3, n (%)	5 (2.6%)	3 (3.6)
5 Minute APGAR	N=192	N=84
Mean (SD)	8.7 (0.8)	8.7 (0.9)
Min, Max	3.0, 10.0	3.0, 9.0
APGAR <3, n (%)	0	0

Source: Table 11-2 Final Report for Study 17P-FU.

The incidence of preterm births in the follow-up population is summarized in Table 19. At each of gestational ages $<37^0$, $<35^0$, and $<32^0$, the percentage of infants in the 17OHP-C treatment groups was numerically lower than that in the placebo group.

Table 19 Pregnancy Outcomes in the follow up Population

Pregnancy Outcome (Weeks Gestation)	17OHP-C N=194 Per cent	Placebo N=84 Per cent
Delivery $<37^0$	30.4%	52.4%
Delivery $<35^0$	14.9%	25.0%
Delivery $<32^0$	7.2%	13.1%

Source: Table 11-2 Final Report for Study 17P-FU.

Division's Comment

- *The 17OHP-C children in the follow-up study may represent a slightly lower risk subset of the total population, as their mean gestational age was one week greater than the total cohort of 17OHP-C children, and they were also more likely to have attained greater gestational age and birthweight than their placebo-exposed peers in the follow-up study.*

6.5.3 Neonatal Morbidity of Enrolled Children

The neonatal morbidities reported at birth for the children enrolled in this study are summarized in Table 20. All occurred with equal or greater frequency in the placebo group as compared to the 17OHP-C group. The differences between the 17OHP-C and placebo groups in the follow-up study were not analyzed statistically. The largest between-group differences (≥4 percentage points) were observed in the incidence of any IVH (1.6% vs. 6.0%) and use of supplemental oxygen (15.5% vs. 21.4%), which were neonatal morbidities that were also lower in the 17OHP-C group in the total population in Study 17P-CT-002.

Table 20 Percentage of Enrolled Neonates Experiencing Morbidities

Morbidity	17OHP-C N=193 (%)	Placebo N=84 (%)
Transient tachypnea	5.2	8.3
Respiratory distress syndrome	9.3	10.7
Bronchopulmonary dysplasia	1.6	3.6
Persistent pulmonary hypertension	0	0
Ventilator support	8.3	10.7
Supplemental oxygen	15.5	21.4
Patent ductus arteriosus	3.1	3.6
Seizures	0	0
Any intraventricular hemorrhage (IVH)	1.6	6.0
Grade 3 or 4 IVH	0.5	0
Other intracranial hemorrhage	0	1.2
Retinopathy of prematurity	2.1	3.6
Proven newborn sepsis	2.1	2.4
Confirmed pneumonia	1.0	2.4
Necrotizing enterocolitis	0	1.2

Source: Table 11-3, Final Report for Study 17P-FU.

The mean and median duration of respiratory therapy for the infants enrolled in the follow-up study were 1.5 and 0.0 days (range: 0.0, 74.0 days) for infants in the 17OHP-C group and 1.9 and 0.0 days (range: 0.0, 44.0 days) for infants in the placebo group.

6.6 Safety Outcomes

Safety assessments were collected via the ASQ, the Survey Questionnaire, and the physical examination. On the Survey Questionnaire, the parent was asked to report any medical diagnosis or operations that occurred between discharge from the birth hospitalization and the time the questionnaire was completed. During the physical examination, the physician was to document any medical abnormality.

Missing data on the ASQ were imputed with the mean of the scores for other items in the same developmental area, as long as ≤ 2 items were missing. If > 2 items were missing, that developmental area was considered missing, and the primary outcome was determined based on the remaining areas. On the PSAI, missing items were imputed with the mean score for that item from the entire sample of same-gender children. If >2 items were missing, the questionnaire was not used. No imputation of missing data was done for other items.

6.6.1 Primary Outcome: Findings from Age and Stages Questionnaire (ASQ)

The ASQ was completed for 275 children, 193 from the 17OHP-C group and 82 from the placebo group. The age of the children at the time of completion of the ASQ ranged from 30 to 64 months; mean age at time of completion did not differ between the 17OHP-C and placebo groups (47.2 vs. 48.0 months). (See Table 21)

Table 21 ASQ – Age of Child at Completion, Source of Information, and Where Completed

	17P N=193[A] n (%)	Placebo N=82[A] n (%)
Age ASQ Completed (months)		
30	1 (0.5)	0
33	9 (4.7)	3 (3.7)
36	30 (15.5)	8 (9.8)
42	49 (25.4)	25 (30.5)
48	32 (16.6)	12 (14.6)
54	38 (19.7)	17 (20.7)
60	34 (17.6)	17 (20.7)
Mean (SD)	47.2 (8.6)	48.0 (8.4)
Median	47.1	48.2
Min, Max	30.2, 63.9	33.5, 64.3
Who Completed Majority of ASQ		
Mother	114 (59.1)	53 (64.6)
Father	2 (1.0)	4 (4.9)
Grandparent	2 (1.0)	0
Foster Parent	1 (0.5)	0
Guardian	2 (1.0)	0
Study Nurse	72 (37.3)	25 (30.5)
Where ASQ Completed		
Home	84 (43.5)	40 (48.8)
Clinical Center	94 (48.7)	34 (41.5)
Home and Clinical Center	15 (7.8)	8 (9.8)

[A] Number of children with ASQ data.
Source: Section 12.3.1, Table 12-1 Final Study 17P-FU-Report

Division Comment

- *At the time that the ASQ was completed, the children in 17OHP-C group tended to be slightly younger, with 21% ≤ 3 years of age, as compared to 14% of placebo children. This might have affected the ability to diagnosis certain developmental problems that may present more noticeably in older children.*

The ASQ was completed predominately by the mother (59.1% 17OHP-C vs. 64.6% placebo) or the study nurse (37.3% vs. 30.5%), and was equally likely to be completed in the home as in the clinical center.

The ASQ responses were categorized to assess communication, gross motor, fine motor, problem solving, and personal-social. Using threshold scores (cutoffs) for normal development, the percentages of children who had scores below the cutoffs for the five areas of development were determined.

Table 22 shows the percentage of children in each treatment group whose ASQ scores suggested developmental problems in at least one of each of the five areas. As the cutoff for identifying a child as needing further developmental evaluation is based, according to the Applicant, on the mean for a normal population, the ASQ would be expected to identify

about 20% of "at risk" children evaluated as possibly delayed. The percentage of children who scored below the cutoff in at least one developmental domain was comparable (27.5% in the 17OHP-C group and 28.0% in the placebo group [p=0.9206]).

The proportion of children below the cutoff in each developmental domain was similar for each treatment group. The area with the highest percentage of children with low scores was fine motor skills, with approximately one in five children scoring below the cutoff (20.7% in the 17OHP-C group vs. 18.3% in the control group). Approximately one in ten children had scores below the cutoff in communication and/or problem solving. Few children had low scores for gross motor and personal-social skills.

Table 22 Percentages of Children in Each Treatment Group Whose ASQ Scores Suggested Developmental Problems

	17OHP-C N=193		Placebo N=82	
	n	%	n	%
Occurrence of score <cutoff on at least one developmental area	53	27.5	23	28.0
Area of Development				
Communication	22	11.4	9	11.0
Gross Motor	5	2.6	3	3.7
Fine Motor	40	20.7	15	18.3
Problem Solving	20	10.4	9	11.0
Personal-Social	7	3.6	1	1.2

Source: Table 12-2, Final Report for Study 17P-FU.

Division's Comment

- *The placebo-exposed children had a greater frequency of very low birthweight (<1500 gm) and delivery prior to 32 weeks (see Table 18 and Table 19). It would be expected that a higher proportion of placebo treated children would be at risk for developmental delays on the basis of these perinatal risk factors. The classification of equal proportions (about 28%) of children in each group as possibly delayed suggests that the 17OHP-C group also resembled an "at risk" group, albeit not as strongly attributable to low birthweight and gestational age. The Applicant did not conduct an analysis adjusting for these risk factors in assessing the proportion of possibly delayed children in each treatment group.*

6.6.2 Secondary Outcomes from Survey Questionnaire

A similar proportion of the children in the 17OHP-C group (99%) and the placebo group (98%) had a completed Survey Questionnaire. Results of the various developmental areas assessed as secondary endpoints are shown in Table 23. There were no marked differences between the groups. A slightly higher proportion of the placebo group had diagnosed problems with motor skills, activity level, communication problems or inability to attend or learn. The most common reported diagnosis was inability to pay attention/learn. When this category is broken down further (not shown in Table 23) the most frequent causes included

"developmental delay," (reported for 2.6% of the 17OHP-C children and 3.7% of the placebo children), and ADHD/ADD, (0.5% in the 17OHP-C group and 2.4% in the placebo group). A child in the 17OHP-C group had a reported diagnosis of mental retardation (Down syndrome) and another child in the 17OHP-C group had a reported diagnosis of autism.

Sensory impairments and need for special equipment were uncommon, but minimally more frequent in placebo children. More than 90% of the children in both treatment groups were reported to have height and weight within the normal range, according to CDC reference growth curves. Almost all of the children in both treatment groups were either in excellent, very good, or good health (98% vs. 95%). No differences in gender-specific roles were noted.

Table 23 Developmental Assessment Based on the Survey Questionnaire

Developmental Area (Scale included in Questionnaire)	Evaluation	17OHP-C N=193		Placebo N=82	
		n	%	n	%
Motor Skills (ECLSK)	% with diagnosis	1[A]	0.5	1[B]	1.2
Activity Level (ECLSK)	% with diagnosis	2	1.0	1	1.2
Communication problems	% with diagnosis	9	4.7	7	8.5
Inability to pay attention/learn	% with diagnosis	8	4.2	5	6.1
Hearing Impairment (NHIS)	% with problem	4	2.1	5	6.1
Vision impairment (NHIS)	% with problem	4	2.1	2	2.4
Need for special equipment	% with problem	1	0.5	1[b]	1.2
Impairment in ability to walk/run/play	% with problem	5	2.6	5	6.1
Overall health	% with "fair health"	4	2.1	4	4.9
	% with "poor health"	0		0	
Height	% below normal	7	3.8	4	5.2
Weight	% below normal	11	5.8	6	7.5
		Mean		Mean	
Gender specific roles (PSAI)	Male score	66.5		67.3	
	Female score	31.8		33.1	

[A] Upper body weakness

[B] Cerebral palsy

Source: Tables 12-5, 12-6, 12-7, 12-8, Final report for Study 17P-FU.

6.6.3 Reported Diagnoses by Health Professionals

Parents/guardians were asked to report for the child any diagnoses made by a health professional at any time between discharge from birth hospitalization and enrollment in the follow-up study. The reported diagnoses are summarized in Table 24. The incidence of each type of reported diagnosis was not meaningfully different (i.e., not > 4 percentage points) between the 2 treatment groups.

Table 24 Reported Diagnoses by Health Professionals

Reported Diagnosis	17OHP-C N=192[A] n (%)	Placebo N=82[A] n (%)
Asthma	39 (20.3)	20 (24.4)
Asthma attack in past 12 months	20 (10.4)	8 (9.8)
Visit to ER or Urgent Care due to asthma in past 12 months	18 (9.4)	7 (8.5)
Eczema or skin allergy	35 (18.2)	12 (14.6)
Ear infections (3 or more)	20 (10.4)	7 (8.5)
Hay fever	19 (9.9)	5 (6.1)
Respiratory allergy	16 (8.3)	9 (11.0)
Developmental delay [B]	14 (7.3)	7 (8.5)
Stuttering or stammering [C]	11 (6.4)	5 (6.6)
Frequent repeated diarrhea or colitis	5 (2.6)	1 (1.2)
Anemia	5 (2.6)	4 (4.9)
Food or digestive allergy	3 (1.6)	3 (3.7)
Seizures or convulsions with fever	3 (1.6)	1 (1.2)
Frequent or severe headaches or migraines [C]	1 (0.6)	2 (2.6)
Diabetes	1 (0.5)	0
Arthritis	1 (0.5)	0
Seizures or convulsions without fever	0	1 (1.2)
Cerebral palsy	0	1 (1.2)
Sickle cell	0	1 (1.2)
Cystic fibrosis	0	0

[A] The number of children for whom the Survey Questionnaire was completed; two children in each treatment group did not have a completed Survey Questionnaire.

[B] Parent/guardian answered "yes" to the question "Has a doctor or other health professional EVER told you that (the child) had any developmental delay?" Per help text provided with the Survey Questionnaire, the parent/guardian was to say "yes" if the health professional diagnosed the child as falling significantly behind age mates in physical, mental, social/emotional, or speech development.

[C] Question answered only for children 3 years or older. Percentages were based on N=171 in 17OHP-C group and N=76 in placebo group.

Source: Table 12-10, Final Report for Study 17P-FU.

6.6.4 Medical Events of Interest

Medical events of interest were potential adverse events that might be attributable to the study drug or to sequelae of prematurity and low birthweight. They were evaluated by integrating data obtained on the ASQ, from the parent on the Survey Questionnaire and from study pediatricians who performed physical exams on the children.

Genital and Reproductive Anomalies

As the study drug involved fetal exposure to a progestin, the occurrence of genital and reproductive anomalies was of particular interest. These were identified by parental reports on the Survey Questionnaire and by physician findings on the physical examination.

Six (3.2%) children in the 17OHP-C group and one (1.2%) child in the placebo group were initially reported by either parent or physician as having genital or reproductive abnormalities. After review of all available data, 2 findings were determined to be

misclassified resulting in genital or reproductive abnormalities in 2.1% (n=4) of the children in the 17OHP-C group and 1.2% (n=1) in the placebo group. The four abnormalities in the 17OHP-C group included:

- micropenis and small scrotal sac noted on study physical examination of a child exposed to 17OHP-C from 19-38 weeks of gestation
- microphallus and Down Syndrome noted on study physical examination of a child exposed from 18-34 weeks of gestation
- surgical correction of undescended testes at an unspecified age in a child exposed from 19-41 weeks of gestation
- early puberty, described by mother as the cause of joint pain that limited the child's ability to walk/run/play, and noted on physical examination (including 4-5 cm breast buds) in a girl exposed from 20-40 weeks of gestation; she was also at the 100th centile for body mass index.

The single genital/reproductive anomaly in the placebo group was described as "sparse public hair" in a 42 month old girl.

Developmental Delays

A second integrated evaluation concerned identification of the "true positives" among those children tagged as potentially at risk for developmental delay based on their ASQ scores. As the purpose of the ASQ is to identify children who may require further evaluation, only some will have confirmation of a developmental delay upon evaluation by a professional. Those children with at least one below-cutoff ASQ score and who also had a parental report of a diagnosis of developmental delay made independently by a professional were reviewed in more detail.

Thirteen (6.7%) of the 193 children in the 17OHP-C group and 8 (9.8%) of the 82 children in the placebo group had an ASQ score below cutoff for at least one developmental area and a reported diagnosis of developmental delay (either in a specific area or overall). The percentages of children evaluated on the ASQ who scored below the cutoff in a specific ASQ developmental area and had at least one reported diagnosis of developmental delay were as follows:

	17OHP-C	Placebo
Communication:	4.7%	8.5%
Gross motor:	1.6%	2.4%
Fine motor:	5.2%	3.6%
Problem solving:	2.6%	6.1%
Personal-social:	2.6%	1.2%

Of the 21 children meeting both criteria, the most common ASQ domains falling below the cutoff were fine-motor and communication for the 17OHP-C group and communication and problem-solving for the placebo children.

Developmental delay, defined as a reported diagnosis by a health professional that the child was falling significantly behind age mates in physical, mental, social/emotional, or speech development, was reported for a similar percentage of children in the 17OHP-C and placebo groups (7.3% vs. 8.5%).

6.6.5 Physical Examination

Physical exams were performed by study physicians on 93% of children in the 17OHP-C group and 87% of the placebo children. Physical examination findings were abstracted from medical records of recent exams for 4% of the 17OHP-C group and 10% of the placebo children; in the remaining cases, no physical findings were available.

Physical findings occurring with disparate distribution over the 2 groups included heart murmurs and irregular rhythm (in ten 17OHP-C and no placebo children), and palpable kidneys (in four 17OHP-C and no placebo children).

6.7 Summary

Study 17P-FU assessed the health status of the children born to women who received weekly intramuscular injections of study drug (17OHP-C or placebo) during Study 17P-CT-002. Only study centers still active in the MFMU Network at the start of Study 17P-FU in the fall of 2004 could participate. Of the 348 infants who were discharged from birth hospitalization at active study sites, 83% (194/234) of the eligible infants in the 17OHP-C group and 74% (84/114) in the placebo group were enrolled in Study 17P-FU. As noted previously, the 17OHP-C children in the follow-up study may represent a slightly lower risk subset of the total population, given their greater mean gestational age as compared to the total cohort of 17OHP-C children, and their greater gestational age and birthweight as compared to their placebo-exposed peers in the follow-up study.

There was no difference between the 17OHP-C and placebo groups in the percentage of children who scored below the cutoff for at least one developmental area of the ASQ. The percentages of children who scored below the ASQ cutoff in each of the individual 5 developmental areas were similar in the 17OHP-C and placebo groups.

Developmental delay, defined as a reported diagnosis by a health professional that the child was falling significantly behind age mates in physical, mental, social/emotional, or speech development, was reported for a comparable percentage of children in the 17OHP-C and placebo groups.

REFERENCES

[1] Mathews TJ, Menacker F, MacDorman MF. Infant mortality statistics from the 2001 period linked birth/infant death data set. Natl Vital Stat Rep 2003; 52:1-28.

[2] Goldenberg RL. The Management of preterm labor. Obstet Gynecol 2002; 100: 1020-37.

[3] Martin JA, Hamilton BE, Sutton PD, Ventura SJ, Menacker F, Munson ML. Births: final data for 2002 Natl Vital Stat Rep 2003; 52:1-113.

[4] Wright VC, Schieve LA, Renolds MA, Jeng G, Kissin D. Assisted reproductive technology surveillance – United States, 2001. MMWR Surveill Summ 2004;53:1-20.

[5] Lockwood CJ, Kuczynski E. Markers of risk for preterm delivery. J Perinat Med 1999;27:5-20.

[6] Iams JD. The epidemiology of preterm birth. Clin Perinatol 2003; 30:651-64.

[7] Berkowitz GS, Papiernik E. Epidemiology of preterm birth. Epidemiol Rev 1993;15:414-43.

[8] Meis PJ, Goldenburg RL, Mercer BM, Iams JD, Moawad AH, Miodovnik M, et al. The preterm prediction study: risk factors for indicated preterm births. Maternal-Fetal Medicine Units Network of the National Institute of Child Health and Human Development. Am J Obstet Gynecol 1998; 178:562-7.

[9] Hogan VK, Njoroge T, Durant TM, Ferre CD. Eliminating disparities in perinatal outcomes – lessons learned. Matern Child Health J 2001;5:135-40.

[10] Collins JW Jr, Hammond NA. Relation of maternal race to the risk of preterm, non-low birth weight infants: a population study. Am J Epidemiol 1996;143:333-7.

[11] Rowley DL. Closing the gap, opening the process: why study social contributors to preterm delivery among black women. Matern Child Health J 2001;5:71-4.

[12] Wadhwa PD, Culhane JF, Rauh V, Barve SS, Hogan V, Sandman CA, et al. Stress, infection and preterm birth: a biobehavioral perspective. Paediatr Perinat Epidemiol 2001;15:17-29.

[13] Meis PJ, Goldenberg RL, Mercer BM, Iams JD, Moawad AH, Miodovnik M, et al. Preterm prediction study: is socioeconomic status a risk factor for bacterial vaginosis in black or in white women? Am J Perinatol 2000;17:41-5.

[14] Alexander GR, Kogan MD, Himes JH, Mor JM, Goldenberg R. Racial differences in birthweight for gestational age and infant mortality in extremely-low-risk US populations. Paediatr Perinat Epidemiol 1999;13:205-17.

[15] Spong CY. No easy answers for predicting preterm labor. Obstet Gynecol 2002;100:219-20.

[16] Iams JD, Goldenberg RL, Mercer BM, Moawad AH, Meis PJ, Das AF, et al. The preterm birth prediction study: can low-risk women destined for spontaneous preterm birth be identified? Am J Obstet Gynecol 2001;184:652-5.

[17] Lockwood CJ. Predicting premature delivery – no easy task. N Engl J Med 2002;346:282-4.

[18] Mercer BM, Goldenberg RL, Moawad AH, Meis PJ, Iams JD, Das AF, et al. The preterm birth prediction study: effect of gestational age and cause of preterm birth on subsequent obstetric outcome. National Institute of Child Health and Development Maternal-Fetal Medicine Units Network. Am J Obstet Gynecol 1999;181:1216-21.

[19] 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 Institutes of Child Health and Development Maternal-Fetal Medicine Units Network. N Engl J Med 1996;334:567-72.

[20] Hibbard JU, Tart M, Moawad AH. Cervical length at 16-22 weeks' gestation and risk for preterm delivery. Obstet Gynecol 2000;96:972-8.

[21] Doland S, Damus K, Green NS, Iams JD, Merkatz IR. Should preterm birth now be classified as a "common complex disorder?" Am J Obstet Gynecol 2004;191:S117.

[22] Romero R, Gomez R, Mazor M, Ghezzi F, Yoon BH. The preterm labor syndrome. In: Elder MG, Romero R, Lamont RF (eds) Preterm Labor, New York: Churchill Livingstone, 1997; 29-49.

[23] Watts DH, Krohn MA, Hillier SL. Eschenbach DA. The association of occult amniotic fluid infection with gestational age and neonatal outcome among women in preterm labor. Obstet Gynecol 1992; 79:351-357.

[24] Romero R, Mazor M. Infection and preterm labor. Clin Obstet Gynecol 1988;31:553-584.

[25] Romero R, Espinoza J, Chaiworapongsa T, Kalache K. Infection and prematurity and the role of preventive strategies. Semin Neonatol 2002; 7:259-274.

[26] Johnson JW, Austin KL, Jones GS, Davis GH, King TM. Efficacy of 17 α-hydroxyprogesterone caproate in the prevention of premature labor. N Engl J Med. 1975; 293 (14):675-80.

[27] Meis PJ, Klebanoff M, Thom E, Dombrowski MP, Sibai B, Moawad AH, et al. Prevention of recurrent preterm delivery by 17 alpha-hydroxyprogesterone caproate. N Engl J Med. 2003; 348 (24):2379-85.

[28] Keirse MJ. Progestogen administration in pregnancy may prevent preterm delivery. Brit J Obstet Gynaecol. 1990;97(2):149-54.