Background Document for Meeting of Advisory Committee for Reproductive Health Drugs
January 20, 2012

NDA 22-139

Progesterone gel (8%)

(Proposed trade name: TBD)

Columbia Laboratories, Inc.

Proposed Indication:

"Progesterone Gel 8% is indicated for the reduction of risk of preterm birth in women with a singleton gestation and a short uterine cervical length in the mid-trimester of pregnancy."

Dosing Regimen:

One applicator (1.125 g) administered vaginally once daily beginning in the second trimester of pregnancy and continuing until 36 completed weeks of gestation or delivery

Prepared by the Division of Reproductive and Urologic Products
Office of New Drugs
Center for Drug Evaluation and Research
Food and Drug Administration

December 22, 2011

The attached package contains background information prepared by the Food and Drug Administration (FDA) for the panel members of the advisory committee. The FDA background package often contains assessments and/or conclusions and recommendations written by individual FDA reviewers. Such conclusions and recommendations do not necessarily represent the final position of the individual reviewers, nor do they necessarily represent the final position of the Review Division or Office. We have brought NDA 22-139 to this Advisory Committee in order to gain the Committee's insights and opinions. The background package may not include all issues relevant to the final regulatory recommendation and instead is intended to focus on issues identified by the Agency for discussion by the Advisory Committee. The FDA will not issue a final determination on the issues at hand until input from the Advisory Committee process has been considered and all reviews have been finalized. The final determination may be affected by issues not discussed at the Advisory Committee meeting.

Table of Contents

Table of Tables

Table of Figures

List of Abbreviations and Definitions

AE	Adverse Event
BMI	Body mass index
BPD	Bronchopulmonary dysplasia
CI	Confidence interval
CMH	Cochran Mantel Haenszel
DDST	Denver Developmental Screening Test
FDA	Food and Drug Administration
IND	Investigational New Drug Application
ITT	Intent to Treat
IUFD	Intrauterine fetal death
IVH	Intraventricular hemorrhage
K-M	Kaplan-Meier
mITT	Modified Intent-To-Treat
NDA	New Drug Application
NEC	Necrotizing enterocolitis
NICU	Neonatal intensive care unit
PK	Pharmacokinetic
PROM	Premature rupture of membranes
PVL	Periventricular leukomalacia
RDS	Respiratory distress syndrome
RR	Relative risk
SAE	Serious adverse event
TVU	Transvaginal ultrasound
US	United States

1. Background

1.1 Objective of Meeting and Overview of Development Program

The purpose of this Advisory Committee meeting is to review and discuss the safety, efficacy, and overall risk/benefit profile of progesterone gel 8%, a progestin indicated for the reduction of risk of preterm birth in women with a singleton gestation and a short uterine cervical length in the mid-trimester of pregnancy.

The primary source of the clinical efficacy and safety data in support of approval of progesterone gel is a single multinational phase 3 clinical trial (Study COL-1620-302, hereafter referred to as Study 302), with limited supportive data provided by post hoc subgroup analysis of an earlier phase 3 trial (Study COL-1620-300, hereafter referred to as Study 300) that had been conducted to support an indication in a different at-risk population.

Study 300 was a prospective, multicenter, randomized, double-blind, placebo-controlled trial, and was the initial trial in the Applicant's clinical development program for progesterone gel for preterm birth. Subjects were selected on the basis of high risk for preterm birth, either because they had a previous spontaneous preterm birth (delivery <35 weeks gestation) or because they had a cervical length of ≤2.5 cm as determined by transvaginal ultrasound (TVU) in the mid-trimester of pregnancy. The primary endpoint was delivery at ≤ 32 0/7 weeks of gestation. As part of Study 300, follow-up data at the ages of six, 12 and 24 months were collected on offspring of women in the trial.

Study 302 was a prospective, multicenter, randomized, double-blind, placebo-controlled trial, conducted in the US and nine countries abroad and was designed to evaluate the safety and the efficacy of progesterone gel 8% in the reduction of preterm birth in women with short cervical length as determined by TVU. The primary endpoint was delivery at ≤ 32 6/7 weeks of gestation. This study was designed following a post hoc subgroup analysis of Study 300, which suggested that in the subset of women with cervical length ≤ 3 cm, progesterone gel showed a trend toward reducing the risk of preterm birth.

Progesterone gel is currently marketed in concentrations of 4% and 8% (NDA 20-701). The approved indications for the progesterone gel 8% are for progesterone supplementation or replacement as part of an Assisted Reproductive Technology ("ART") treatment for infertile women with progesterone deficiency, and for the treatment of secondary amenorrhea in women who have failed to respond to treatment with [progesterone gel] 4%." Progesterone gel has not been approved in any country for reduction of the risk of preterm birth indication.

1.2 Issues for Committee Consideration

Committee members are asked to reflect upon the following issues for consideration as they review the information provided in this Background Document.

Issues for consideration include the following:

- Has the Applicant provided sufficient information to conclude that progesterone gel reduces the risk of preterm birth in women with a singleton gestation and a short uterine cervical length at mid-trimester of pregnancy, given that statistically significant efficacy was not demonstrated in US subjects?

- Do you believe that there is any explanation, based on the data provided in the NDA, for the difference in efficacy results in the US and foreign populations? If yes, do you believe that the explanation could be adequately addressed in labeling so that progesterone gel could be used safely and effectively in the US population?
- Has the Applicant provided sufficient information to conclude that the safety profile for progesterone gel is acceptable for the proposed indication?
- Is the overall risk/benefit profile of progesterone gel acceptable to support approval of this product in the US for the proposed indication?
- If not, do you have recommendations as to how efficacy and/or safety could be investigated further in the US population (e.g., a new study)?

1.3 Prevention of Preterm Birth

Preterm birth, defined as delivery prior to 37 completed weeks of gestation, is a significant public health problem in the US and globally. In the US, the prevalence of preterm birth has increased since 1990, currently affecting 12.8% of all births and 11% of singleton births[1]. Although there are a number of diagnostic tests proposed to identify women at risk for preterm labor and several medications are used off-label in an attempt to stop preterm labor, there are no data indicating a benefit on neonatal morbidity or mortality from any of these interventions. In February 2010, the first drug for the reduction of the risk of preterm birth was approved by FDA. This drug, Makena, or hydroxyprogesterone caproate, is indicated "to reduce the risk of preterm birth in women with a singleton pregnancy who have a history of singleton spontaneous preterm birth." This is not the same population for which the product currently under review would be indicated, if approved; nevertheless it is likely that there would be some overlap between the two populations.

Progesterone gel, the focus of this Advisory Committee meeting, if approved, would be the first and only product approved for reduction of risk of preterm birth in women with a singleton pregnancy and a short cervix at the mid-trimester of pregnancy. A short cervix is a recognized risk factor for or precursor to preterm delivery. While the relationship is likely to be continuous, the risk of preterm birth has been demonstrated to increase markedly when the cervix is ≤ 3 cm at 24 weeks of gestation, and a clinically important short cervix is generally defined as a cervix measured by vaginal sonography as ≤ 2.5 cm in length. This corresponded to the 10^{th} centile in a large prospective study, and was associated with a relative risk of 6.2 for preterm birth [2].

1.4 Clinical Pharmacology of Progesterone Gel

Progesterone is present in very high quantities during pregnancy and the concentration increases as gestation progresses. The specific role, if any, of progesterone with respect to preventing or initiating parturition has not been characterized in humans. The Applicant studied the serum levels of progesterone pre- and post-dose in a small number of pregnant women taking progesterone gel (see Section 2.3).

The current product, progesterone vaginal gel 8% (w/w), is a sustained-release formulation for vaginal application and contains 90 mg/dose in a polycarbophil-based gel. Women are instructed to administer the drug once daily, from the second trimester of pregnancy until the earlier of completion of 36 weeks of gestation or delivery.

1.5 Regulatory Guidance for the Development of Progesterone Gel

The Applicant conducted the drug development program for this indication under IND 68,884. The IND was opened with a protocol for the phase 3 Study 300 in December, 2003. The Division of Reproductive and Urologic Products (also referred to as the Division) provided guidance to the Applicant on the protocol for Study 300 in February 2004. The most important points conveyed to the Applicant were:

- The proposed endpoint of delivery at ≤ 37 weeks gestation was of limited clinical importance; the Division requested use of the endpoint of delivery at ≤ 32 weeks gestation
- Pharmacokinetic (PK) data was requested in a subset of subjects
- Autopsy data and other relevant information (such as placental histopathology for all stillbirths) was requested for all stillbirths and infant deaths
- Follow-up information on infants up to 2 years of age was requested
- While in general, two adequate and well-controlled trials are required for approval a drug product for a new indication, one large phase 3 trial might be adequate IF it (1) showed a robust, statistically significant reduction in preterm births at ≤ 32 weeks gestation with no suggestion of an increase in intrauterine deaths or neonatal morbidity or mortality and (2) included adequate infant follow-up

During a teleconference in April 2004, the Division indicated that the dose selection was acceptable. It was requested that fetal and placental pathology specimens be evaluated by a pathologist with expertise in perinatal pathology. In the infant follow-up program, the Division requested that all abnormal findings be evaluated by appropriate pediatric specialists. The Division provided further clarification of a "compelling and robust reduction in preterm births" as a p-value of <0.01, no increase in morbidity/mortality of fetuses, neonates, infants, or mothers, and outcome endpoints of clinical benefit (such as shortened hospital stay, reduced neonatal morbidity). If planning to rely upon a single trial for approval, the Applicant was encouraged to include as a co-primary endpoint reduction in the incidence of significant clinical morbidity.

An additional guidance meeting was held in July 2007 to discuss a revised indication sought following failure of Study 300 on its primary endpoint of reduction of preterm birth in women with a history of preterm birth. The study had included a small subset of women with a shortened cervix, and post hoc analysis of data from this subset indicated that progesterone gel might be of benefit in this population. The Applicant initially asked whether, based on the post hoc analysis, the NDA could be submitted for review. The Division informed the Applicant that approval of the drug product would require at least one additional adequate and well-controlled study in a prospectively defined population at risk of preterm birth, using a primary endpoint of delivery at ≤ 32 weeks gestation. Demonstration of benefit in actual clinical outcomes for the infants would also be needed, such as significant improvement on a composite clinical endpoint measuring neonatal morbidity/mortality.

The Applicant discussed conduct of an additional phase 3 study in women with a short cervix, and initially proposed defining "short cervix" as < 2.8 cm. The Division stated that the literature generally defines short cervix as < 2.5 cm, but that a robust finding in a population at any prespecified shortened cervical length would be of clinical importance.

The Division also agreed that a reduction in preterm birth ≤ 32 weeks gestation could be a surrogate endpoint for infant morbidity, as defined in CFR 314.50. The Applicant was encouraged to add a composite neonatal morbidity/mortality measure to their clinical trial that included any liveborn infant who experienced death, respiratory distress syndrome (RDS), bronchopulmonary dysplasia (BPD), Grade 3 or 4 intraventricular hemorrhage (IVH), proven sepsis or necrotizing enterocolitis (NEC). The Division concurred that infant follow-up data on the cohort of offspring from Study 300 would likely be sufficient to support submission of an NDA, but that additional safety data might be required in the postmarketing period.

Another teleconference was held in April 2010 to discuss the Statistical Analysis Plan for Study 302. The primary efficacy endpoint for this study was delivery at ≤ 32 6/7 weeks (i.e., < 33 weeks) gestation, with a supportive endpoint relating to neonatal morbidity/mortality. The following points were discussed:

- The Intent to Treat (ITT) population, defined as subjects who received at least one dose of study drug, should include all treated subjects regardless of whether or not they had a recorded delivery date; subjects should be censored at the last date they were known to be pregnant, rather than excluded from the ITT. For subjects whose last date of contact was prior to 33 weeks, those randomized to active treatment should be considered as failures, while those randomized to placebo should be considered successes, to provide a "worst case" analysis.
- The primary efficacy analyses should estimate and evaluate risk differences, not odds ratios
- The procedure used to pool small study sites should be described
- The neonatal mortality/morbidity index will have an important impact on the approvability of the drug; the protocol should specify definitions for each morbid event
- Neonatal mortality should be evaluated individually, while the total morbidity score should be based on the number of morbid events occurring in each neonate; BPD should be added to the composite score
- The Applicant proposed a 0-4 point score for morbidity/mortality, in which 0-2 represents the equivalent number of morbid events, 3 includes 3-5 morbid events and 4 includes neonatal death. The Division also requested a sensitivity analysis using a 0-6 point scale.

A pre-NDA meeting was held in February 2011. The Applicant was asked to provide developmental follow-up on offspring born to mothers enrolled in Study 300; the same follow-up should be done on offspring from Study 302, but this could be done as a postmarketing requirement if the product were approved. The Applicant was also asked to provide a literature review on the safety of progesterone for all indications of progesterone in pregnancy. The Applicant was asked to provide Kaplan-Meier analyses using time from randomization to delivery (to evaluate duration of treatment), and to ensure that analyses based on gestational age at delivery account for left truncation resulting from enrollment at different gestational ages. The Statistical Analysis Plan was otherwise acceptable.

2. Clinical Development of Progesterone Gel

2.1 Overview of Product Development

The development program for progesterone gel for the preterm birth indication consisted of one phase 1 bioavailability study, two phase 3 randomized clinical trials in slightly different at-risk populations of pregnant women, and a follow-up study of offspring of exposed women in the first phase 3 study.

2.2 Overview of Pharmacology and Toxicology

No new animal studies were performed in support of this NDA; the pharmacology/toxicology information was cross-referenced to that contained in NDA 20-701 and updated through a literature search.

2.3 Overview of Clinical Pharmacology

The product used in the clinical and clinical pharmacology studies was identical to the to-be-marketed product for the proposed indication (as well as to that currently marketed for gynecologic indications). A single phase 1 study, Study COL-1620-301, was conducted to evaluate the pre- and post-dose serum concentrations of progesterone in pregnant women who received progesterone gel 8% at 28 weeks of gestation. Subjects at risk of preterm birth by virtue of having had a previous preterm birth were enrolled and started on daily treatment of progesterone gel between 18 0/7 and 22 6/7 weeks of gestation. Serum levels of progesterone were drawn pre- and post-dose at Week 28 and post-dose at Week 36.

All serum progesterone concentrations were within the normal reference range. At Week 28, pre-dose levels (N=19) ranged from 57 to 175 ng/ml. Post-dose levels (N=21) ranged from 80 to 211 ng/ml, with levels generally increasing from pre- to post-dose (the range of change was -14 to 58 ng/ml; median change was 17 ng/ml). At Week 36 (N=14), post-dose levels ranged from 115 to 398 ng/ml.

Division Comment

Individual progesterone levels varied quite widely, and there are known to be diurnal variations in progesterone levels in pregnancy; therefore, it is difficult to evaluate the impact of exogenous progesterone on the serum levels obtained in this study.

2.3.1 Pharmacodynamics / Mechanism of Action

The mechanism of action of progesterone gel with respect to risk of preterm birth is unknown.

2.4 Overview of Clinical Studies

The clinical portion of the NDA includes three studies: Study COL-1620-301 was a PK study of progesterone dosing in pregnancy, Study COL-1620-300 was a phase 3 study in women at risk for preterm birth on the basis of a prior preterm birth (this study also contained the infant follow-up data), and Study COL-1620-302 was a phase 3 study in women with a short cervix in the mid-trimester of pregnancy.

2.4.1 Phase 3 Clinical Studies

The Applicant completed two phase 3 studies. As noted in Section 1.1, Study 300 was the initial trial, intended to evaluate progesterone gel in a population of women at high risk of preterm birth either because they had a history of a previous spontaneous preterm birth (delivery <35 weeks gestation) or because they had a short cervix (≤ 2.5 cm) measured by

TVU in the mid-trimester of pregnancy. The overall results of this trial did not demonstrate efficacy of progesterone gel in the population of women at risk because of a previous preterm birth (p = 0.69 for the primary endpoint). However, the Applicant conducted a post hoc analysis limited to the subgroup of women with the short cervix risk factor. Based on favorable results in this subgroup, the Applicant designed Study 302 to evaluate efficacy in a population at risk solely on the basis of short cervical length. The major focus of the Division's efficacy evaluation is on the results of Study 302; the subgroup analysis of Study 300 will be briefly discussed.

2.5 Basis for Dose Selection

The Applicant studied only a single dose, 90 mg. It is not stated how this dose was selected; however, the Applicant notes that a previous randomized clinical trial of 100 mg progesterone, administered by vaginal suppository to women at high risk of preterm birth, demonstrated a lower preterm delivery rate at 32, 34, 35, and 37 weeks of gestation.

3. Objectives and Design of Phase 3 Trials

Study 302 provided the primary data to support the efficacy and safety of progesterone gel for reduction of the risk of preterm birth in women with a shortened cervix. Study 300 enrolled women with different risk factors for preterm birth (a prior history of spontaneous preterm birth) as well as a subset of women with a short cervix. Data from this subset, which met different enrollment criteria relating to cervical length compared to those used in Study 302, provided limited supportive evidence relating to efficacy.

An overview of the two studies is provided in Table 1.

Table 1 Phase 3 Studies for Progesterone Gel

Study Number (No. of Sites / Country) Dates of Study Conduct	Subject Population	Primary Endpoints	Treatments	R (ITT)[1]	Design
COL-1620-302 (21 / Non-US, 23 / U.S.) Mar. 2008 to Oct. 2010	Women 15 to 45 years of age, cervical length of 1.0 to 2.0 cm; between 19 0/7 to 23 6/7 weeks gestation	Frequency of preterm birth of ≤ 32 6/7 weeks gestation	Progesterone Placebo **Total**	236 (235) 229 (224) **465 (459)**	Placebo-controlled, randomized, double-blind, multicenter, multinational parallel group
COL-1620-300 (14 / Non-US, 39 / U.S.) Mar. 2004 to Jan 2007	Women 15 to 45 years of age, history of preterm birth ***OR*** cervical length of 2.5 cm or less; between 16 0/7 to 22 6/7 weeks gestation.	Frequency of preterm birth of ≤ 32 0/7 weeks gestation	Progesterone Placebo **Total**	336 (321) 332 (316) **668 (637)**	

[1]R = Randomized Subjects, ITT= Intend To Treat

3.1 Study 302

3.1.1 Study Objectives

Primary Objective

The primary objective of the study was to evaluate the reduction in the frequency of preterm birth (delivery) at ≤ 32 6/7 weeks gestation.

Secondary Objectives

The secondary objectives were to demonstrate:

1. Reduction in the frequency of perinatal mortality and neonatal morbidities including RDS, BPD, grade III or IV IVH, periventricular leukomalacia (PVL), proven sepsis, NEC and neonatal intensive care unit (NICU) days as assessed individually and through a composite score;
2. Reduction in the frequency of preterm birth (delivery) at ≤ 27 6/7, ≤ 34 6/7, and <36 6/7 weeks gestation;
3. Reduction in the frequency of admission for preterm labor;
4. Assessment of the admission-to-delivery interval in subjects receiving tocolytic therapy for preterm labor;
5. Assessment of APGAR (activity, pulse, grimace, appearance, respiration) scores, length, weight, and head circumference at birth, and incidence of congenital anomalies;
6. Assessment of other indicators of neonatal morbidity such as admission to the NICU, the duration of stay in the NICU, and the total hospital stay.

Division Comments:

- **PVL was not one of the neonatal morbidities agreed upon by the Division and was not specified in the protocol. However, it was added in the Statistical Analysis Plan.**
- **The protocol specified "neonatal" mortality as a component of this composite score, but the Statistical Analysis Plan clarified that this category also includes IUFDs and stillbirths, and that this had always been the intention. The Applicant considered "perinatal" mortality to be a more accurate descriptor.**

3.1.2 Overall Study Design and Conduct

Study 302 was a randomized, double-blind, placebo-controlled, multicenter, multinational study in women with a cervical length of 1-2 cm. The trial was conducted at 23 sites in the US and 21 sites in nine foreign countries (2 sites in Belarus, 2 sites in Chile, 1 site in Czech Republic, 5 sites in India, 3 sites in Israel, 1 site in Italy, 1 site in Russia, 1 site in South Africa, and 5 sites in Ukraine).

Study Schedule and Conduct

Pregnant women aged 15 to 45 who were between 19 0/7 and 23 6/7 and who had a cervical length measured by TVU of 1-2 cm were eligible for screening. Randomization was stratified by study site and by risk strata (defined as presence or absence of a prior spontaneous preterm birth). Once randomized, subjects began taking study drug between 20 0/7 and 23 6/7 weeks of gestation, depending on when they enrolled, and continued dosing until the earlier of 36 6/7 weeks or delivery. Study drug was stopped if women had premature rupture of membranes. Table 2 provides the Schedule of Events.

Table 2 Study 302: Schedule of Events

Assessments	Study Period/Visit			
	Enrollment and Randomization	Biweekly Visits	Delivery or Maternal Completion	Post-delivery
Informed consent	X			
Inc/Exc criteria	X			
Demographics	X			
Obstetrics history	X			
Medical history	X			
Current medications	X			
Vital signs	X			
Height and weight	X			
Physical examination	X			
TVU and TAU	X			
Medical release form	X			
Randomization	X			
Dispense test article	X	X		
Test article treatment	X	X		
AE assessment		X	X	
Brief physical/obstet exam		X		
Fundal height	X	X		
Fetal heart rate	X	X		
Concomitant medications		X	X	
Treatment compliance		X	X	
Labor/Delivery timing			X	
Health of mother			X	
Health of infant			X	
Fetal/Infant mortality/morbidity		X	X	X

AE=adverse event; Inc/Exc=inclusion/exclusion; obstet exam= obstetrical examination; TAU=transabdominal ultrasound; TVU=transvaginal ultrasound.

Source: Applicant's Study Report for Study 302, Table 9-1

3.1.3 Eligibility Criteria

Significant inclusion criteria included women aged 15 to 45 years who were pregnant with a singleton gestation, were between 19 0/7 and 23 6/7 weeks of gestation, and who had a cervical length measured by TVU of 1-2 cm. Important exclusion criteria included planned cervical cerclage, any of a number of medical conditions, including seizure disorder, chronic hypertension, unstable psychiatric disorder, uncontrolled diabetes, chronic renal failure, active or history of thromboembolic disorder, liver dysfunction or genital or breast malignancy. Pregnancy complications that precluded entry included fetal anomalies or chromosomal abnormality, preterm rupture of membranes, vaginal bleeding, known or suspected chorioamnionitis, signs of preterm labor at enrollment, and placenta previa.

3.1.4 Efficacy Assessments

3.1.4.1 Analysis Populations

Efficacy was evaluated using the following analysis populations:

1. **Intent-To-Treat (ITT)** analysis set consisted of all subjects who were randomized and received at least one dose of study drug.
2. **Modified Intent-To-Treat (mITT)** analysis set consisted of all ITT subjects who had ≥ 80% treatment compliance during the specified time treatment (i.e., from Week 20 0/7 to 36 6/7 or delivery or premature rupture of membranes), a documented delivery date, and no cerclage performed.

   Treatment compliance was calculated according to the following formula:

$$\frac{\text{Number of vaginal applicators used since last visit x 100}}{\text{Number of vaginal applicators that should have been used since last visit}}$$

The Applicant considered the ITT population to be the primary analysis population for efficacy.

Division Comment

The mITT analysis set is similar to a Per Protocol analysis set, and was given little weight by the Division.

3.1.4.2 Efficacy Endpoints and Analyses

Primary Endpoint

The primary efficacy endpoint was the proportion of subjects who delivered at a gestational age of ≤ 32 6/7 weeks. Gestational age was based on the participant's reported last menstrual period and fetal sonographic biometry, and was calculated as [40 + (delivery date - accepted EDC)/7].

The primary analysis was performed using a Cochran-Mantel-Haenszel (CMH) test stratified by primary pooled study site and risk strata, which was defined as the presence or absence of a previous preterm birth at 20-35 weeks gestation in the prior pregnancy. The p-value from this analysis was assessed at the two-sided significance level of 0.05. The primary efficacy analysis was based on the ITT analysis set, with the mITT analysis set used for a secondary analysis. Risk difference and relative risk (RR) estimates along with 95% CIs were constructed. These estimates were generated overall and by risk strata. In order to conclude that there was statistically significant evidence of a reduction in the risk for preterm birth at ≤ 32 6/7 weeks associated with progesterone gel, it was necessary that (1) the 95% CI for the risk difference not cross 0 and the direction of the difference in risk favored progesterone gel and (2) the CI for the RR did not cross unity.

Additional statistical analyses were conducted using a logistic regression model with terms for primary pooled study site, risk strata, treatment group and covariates consisting of gestational age at entry in the study, maternal age, cervical length, body mass index (BMI), and race. The primary pooled study site by treatment interaction effect using the logistic regression model was also tested at the two-sided significance level of 0.10.

Gestational age at delivery was presented graphically using Kaplan-Meier (K-M) estimates for each treatment group. Differences between the treatment groups were tested using a log-

rank test stratified by primary pooled study site and risk strata. The time from randomization to delivery was included as a K-M estimated curve. In addition, K-M estimated survival curves for gestational age at delivery and for gestational age at perinatal mortality were generated adjusting for left truncation based on gestational age at first dose using a null Cox proportional hazards model.

The Applicant pooled sites based on geographic location and enrollment at each study site. Pooling by region (US vs. non-US) was one of three planned methods of pooling.

For subjects with missing information on gestational age at delivery, a "worst case" imputation plan was followed, in which placebo subjects were considered to have had term deliveries, while progesterone gel subjects had the last date of contact imputed as their delivery date. All other data handling was done on an observed case basis, with no other imputation.

There were two planned subgroup analyses: a previous preterm birth (≤ 35 0/7 weeks) vs. no previous preterm birth, and cervical length ≤ 1.7 cm vs. > 1.7 cm. Over the course of the review, FDA requested a number of additional subgroup analyses. No interim analyses were planned or conducted.

Secondary Endpoints

The key secondary efficacy endpoint was the proportion of subjects with a composite score for perinatal mortality and neonatal morbidities that encompassed RDS, BPD, grade III or IV IVH, PVL, proven sepsis, or NEC. The primary composite score as specified in the Statistical Analysis Plan was a 0-4 scale, which was derived as an ordinal severity scale scored as:

- 0 - no events
- 1 - one morbidity event from among RDS, BPD, IVH, PVL, proven sepsis, or NEC, and no perinatal mortality
- 2 - two morbidity events and no perinatal mortality
- 3 - three or more morbidity events and no perinatal mortality
- 4 - perinatal mortality

The Applicant also evaluated a 0-6 composite score, which had the same definitions as the 0-4 scale for 0 to 2, and defined 3 as three morbid events and no perinatal mortality, 4 as four morbid events and no perinatal mortality, 5 as five morbid events and no perinatal mortality, and 6 as perinatal mortality. The Applicant also evaluated modified 0-4 and 0-6 composite scores that included number of days in the NICU as part of the score definitions.

The stratified CMH test as described above for the primary efficacy variable was also used for the composite mortality/morbidity score.

Division Comment

The Division initially had requested the Applicant to use the 0-6 composite score as the primary composite score, but had agreed that the 0-4 score could be used as long as the 0-6 score was included as a sensitivity analysis. The Division also had specifically asked that NICU days be excluded from the score definitions because NICU length of stay was believed to be a surrogate measure for serious health problems that would be captured by the morbidity categories. In addition, the Division had concerns about practice differences over the different regions participating in the study; in particular, availability of NICU care and criteria for NICU admission were anticipated to be highly variable across the international sites.

The other secondary efficacy endpoints were the proportion of subjects who delivered at ≤ 27 6/7 weeks, ≤ 34 6/7 weeks, ≤ 36 6/7 weeks, and term delivery (≥ 37 weeks), as well as the actual gestational age at delivery. Analyses using the CMH test and logistic regression model for primary efficacy endpoint were performed for the secondary endpoints.

3.2 Study 300

3.2.1 Study Objectives

The primary objective of this study was to demonstrate a reduction in the frequency of delivery at ≤ 32 0/7 weeks of gestation in pregnant women who had a previous preterm birth.

The secondary objective was to evaluate the frequency of significant infant morbidity in the initial hospitalization following birth.

3.2.2 Overall Study Design and Conduct

Study 300 was conducted at 53 study sites in six countries (1 in El Salvador, 2 in Chile, 2 in Czech Republic, 6 in India, 3 in South Africa and 39 in the US).

Screening of potential study participants was conducted when their pregnancies were between 16 0/7 and 22 6/7 weeks of gestation. Subjects could qualify for the trial on the basis of two possible risk factors for previous preterm birth – a prior spontaneous preterm birth (planned to make up about 95% of the population) or a short cervix (≤ 2.5 cm, planned to make up about 5% of the study population). Subjects who had had a spontaneous preterm delivery in the previous pregnancy were randomized between 18 0/7 to 22 6/7 weeks. For subjects whose only risk factor for preterm birth was a short cervix, randomization occurred from 20 0/7 to 22 6/7 weeks. Once randomized, the subjects began treatment with the allocated study drug and administered it daily until 37 0/7 weeks gestational age, development of preterm rupture of membranes, or delivery, whichever occurred first. Subjects were randomized in a 1:1 ratio to either progesterone gel or a matching placebo gel. Thus, this trial consisted of two distinct clinical populations with separate randomization schemes and slightly different gestational ages at which treatment could be initiated.

3.2.3 Eligibility Criteria

Subjects were pregnant women between 18 and 45 years with a pregnancy between 16 0/7 to 22 6/7 weeks of gestation. Women could qualify on the basis of a documented spontaneous singleton preterm birth occurring between 20 0/7 to 35 0/7 weeks of gestation in the immediately prior pregnancy, or on the basis of a cervical length ≤ 2.5 cm measured by TVU. All subjects were screened by TVU at baseline, and those who were potential candidates for cerclage placement according to local practice standards were excluded. Women were also excluded if they had any of a number of medical conditions, including seizure disorder, chronic hypertension, unstable psychiatric disorder, uncontrolled diabetes, chronic renal failure, active or history of thromboembolic disorder, liver dysfunction or genital or breast malignancy. Pregnancy complications that precluded entry included fetal anomalies or chromosomal abnormality, preterm rupture of membranes, vaginal bleeding, known or suspected chorioamnionitis, signs of preterm labor at enrollment, and placenta previa or low-lying placenta.

3.2.4 Efficacy Assessments

3.2.4.1 Analysis Populations

The Applicant planned to enroll about 95% of subjects on the basis of a previous preterm birth, and 5% on the basis of a short cervix. The ITT population included only those women with a previous preterm birth; there was also a small "SHCX" population that included those women with cervical length ≤ 2.5 cm.

Division Comment

The Applicant describes the SHCX subgroup as "a pilot study in a new population."

3.2.4.2 Efficacy Endpoints and Analyses

Primary Endpoint

The primary efficacy endpoint was the proportion of women in the ITT population (women with previous preterm birth) who delivered at ≤ 32 0/7 weeks of gestation.

The primary efficacy analysis used an overall test of the proportions in each treatment arm, including 95% CI. Supportive analyses on this endpoint were also done using a logistic regression model with the following covariates:

- Gestational age of previous preterm birth
- Gestational age at first dose
- Maternal age
- Pretreatment cervical length
- BMI
- Race
- Investigator effect

An investigator by treatment interaction was also evaluated.

Secondary Endpoint

The main secondary efficacy endpoint was the frequency of significant infant morbidity in the birth hospitalization. Other secondary endpoints looked at the proportion of deliveries occurring at ≤ 28 0/7, ≤ 35 0/7, and ≤ 36 6/7 weeks of gestation. An additional secondary outcome was the proportion of subjects in the SHCX group (i.e., those whose only risk factor was a short cervix) who delivered at ≤ 32 0/7 weeks of gestation. No adjustments were made to account for multiple comparisons.

Post Hoc Analyses

Only nine subjects were enrolled in the original short cervix subgroup of Study 300 (i.e., with cervical length < 2.5 cm): four in the progesterone gel group and five in the placebo group. However, the Applicant conducted a post hoc analysis in which subjects were divided into quartiles of cervical length, and the primary endpoint was evaluated by treatment assignment in each quartile.

Division Comment

At the July 2007 meeting, the Division informed the Applicant that it considered this post hoc analysis to be hypothesis-generating only, and that an additional study to test this hypothesis would be needed if the Applicant wished to pursue an indication for treatment of a population in which short cervix was the risk factor for preterm birth.

Additional Post Hoc Efficacy Analyses

The Applicant also conducted secondary analyses of the treatment effect in women with cervical length ≤ 2.8 cm (there were 28 women in the progesterone gel treatment arm and 30 in the placebo arm in this subset) and those with cervical length ≤ 3 cm (there were 58 women in each arm in this subset). The Applicant reportedly selected the 3 cm cutpoint on the basis of the quartile analysis, and the 2.8 cutpoint because this cervical length correlated with improvement in infant outcomes.

4. Efficacy Findings

4.1 Study 302

4.1.1 Subject Enrollment and Disposition

A total of 465 pregnant women were randomized in a 1:1 ratio at 44 sites in 10 countries. Of these randomized subjects, five placebo subjects and one progesterone gel subject did not receive any treatment. Among the 459 treated subjects, 445 (98% of those on active drug and 96% on placebo) completed the study. Details of subject disposition are summarized in Table 3.

Table 3 Study 302: Disposition of Subjects (ITT Population)

Category	Placebo		Progesterone Gel		Total	
	N	%	N	%	N	%
Randomized and Treated (ITT)	224	100	235	100	459	100
Completed the Study	215	96.0	230	97.9	445	96.9
Discontinued Study Drug	9	4.0	5	2.1	14	3.1
Reason for Discontinuation:						
Adverse Event: Stillbirth	3	1.3	3	1.3	6	1.3
Adverse Event: IUFD	2	0.9	2	0.9	4	0.9
Lost to Follow-up	3	1.3	0	0	3	0.7
Other	1	0.4	0	0	1	0.2

Percentage based on the total number of treated subjects within each corresponding treatment group
IUFD = intrauterine fetal death
Source: FDA Statistical Analysis and Review

The details of subject disposition by US vs. non-US region are presented in Table 4. Overall, the distribution was fairly equal, with 45% of the subjects enrolled in US sites compared to 55% of subjects enrolled in non-US sites. However, the compliance rate in the US region was much lower than that in the non-US region (US region: 22.8%; non-US: 2.4%), so the mITT analysis set included only 72% of all US ITT subjects compared to 94% of non-US ITT subjects.

Table 4 Study 302: Disposition of Subjects by Region (ITT Population)

Category	Placebo		Progesterone Gel		Total	
	N	%	N	%	N	%
US Sites						
Randomized and Treated (ITT) [1]	99	44.2	107	45.5	206	44.9
Discontinued Study Drug [1]	1	0.4	3	1.3	4	0.9
Non-compliant [2]	17	17.2	30	28.0	47	22.8
mITT Analysis Set [2]	77	77.8	72	67.3	149	72.3
Non-US Sites						
Randomized and Treated (ITT) [1]	125	55.8	128	54.5	253	55.1
Discontinued Study Drug [1]	8	3.6	2	0.9	10	2.2
Non-compliant [2]	4	3.2	2	1.6	6	2.4
mITT Analysis Set [2]	116	92.8	122	95.3	238	94.1
All ITT Analysis Set	224	100	235	100	459	100

[1] Percentage based on all ITT subjects within each corresponding treatment group.
[2] Percentage based on ITT subjects within each region and corresponding treatment group.
Source: FDA Statistical Analysis and Review

Division Comments

- **The rate of premature discontinuation was higher in the US sites, and slightly higher in the progesterone gel subjects, while in the non-US regions, placebo subjects were more likely to discontinue early.**
- **The pattern of noncompliance also varied over US and non-US regions: in the US, subjects randomized to progesterone gel were more likely to be noncompliant than placebo subjects; this tendency was reversed in the non-US region.**

4.1.2 Demographic and Baseline Characteristics

Demographic and baseline characteristics for all treated subjects by region are presented in Table 5. Although demographic and baseline characteristics appear to be balanced across treatment groups, there are differences by region in age, race, BMI and cervical length.

The majority of subjects in the US were Black (60%) and Caucasian (29%), while the majority of subjects in non-US sites were Asian (59%) and Caucasian (23%). The distribution of race by region is shown in Table 6. The contribution of each region to the overall ethnic distribution was unbalanced. The vast majority of Asian subjects were in the non-US region, while the majority of Black and Other subjects were from the US region. Caucasians were distributed proportionally in each region.

The mean BMI in US subjects was 29.2 kg/m^2, while the mean BMI in non-US subjects was 22.4 kg/m^2. In addition, the US subjects were older and had shorter cervical length compared to non-US subjects.

Division Comments

These variables that are discrepant between US and non-US regions are risk factors for preterm birth. However, it is not easy to conclude that one region was at higher risk than the other, because certain risk factors (e.g., low maternal weight) were more common in the non-US regions, while others (e.g., shorter cervical length) were more frequently observed in the US region.

Table 5 Study 302: Demographics and Baseline Characteristics (ITT Population)

Characteristic Mean (SD)	Placebo			Progesterone Gel			Overall		
	US N=99	Non-US N=105	Total N=224	US N=107	Non-US N=128	Total N=235	US N=206	Non-US N=253	Total N=459
Age (years)	27.2 (5.8)	25.3 (4.3)	26.1 (5.1)	26.9 (6.3)	26.2 (5.5)	26.5 (5.8)	27.0 (6.0)	25.8 (4.9)	26.3 (5.5)
Race n(%)									
Asian	1 (1%)	76 (61%)	77 (34%)	3 (3%)	73 (57%)	76 (32%)	4 (2%)	149 (59%)	153 (33%)
Black	58 (59%)	9 (7%)	67 (30%)	66 (62%)	10 (8%)	76 (32%)	124 (60%)	19 (8%)	143 (31%)
Caucasian	31 (31%)	38 (30%)	69 (31%)	28 (26%)	45 (35%)	73 (31%)	59 (29%)	83 (33%)	142 (31%)
Other	9 (9%)	2 (2%)	11 (5%)	10 (9%)	0	10 (4%)	19 (9%)	2 (<1%)	21 (5%)
BMI	29.1 (6.7)	22.2 (4.9)	25.2 (6.7)	29.3 (5.8)	22.7 (5.0)	25.6 (6.3)	29.2 (6.3)	22.4 (4.9)	25.4 (6.5)
Cervical Length (cm)	1.61 (0.32)	1.77 (0.23)	1.70 (0.29)	1.64 (0.29)	1.79 (0.20)	1.72 (0.25)	1.62 (0.30)	1.78 (0.22)	1.71 (0.27)
Gestational Age at First Dose (Weeks)	22.0 (1.3)	21.6 (1.4)	21.8 (1.4)	21.9 (1.4)	21.8 (1.3)	21.9 (1.4)	22.0 (1.4)	21.7 (1.4)	21.8 (1.39)

Source: FDA Statistical Analysis and Review

Table 6 Study 302: Distribution of Race by Region (ITT Population)

Race	US N=206		Non-US N=253		Total N=459
	n	%	n	%	N
Asian	4	3	149	97	153
Black	124	87	19	13	143
Caucasian	59	42	83	58	142
Other	19	90	2	10	21

Source: FDA Statistical Analysis and Review

4.1.3 Efficacy Findings

4.1.3.1 Primary Efficacy Endpoint and Analysis

The results of the Applicant's primary efficacy analysis based on the ITT population (and analysis of treatment effect on deliveries at other weeks of gestation) are shown in Table 7. Four subjects in the placebo group did not have a documented delivery date and were counted as term deliveries at 37 weeks. All subjects in the progesterone gel group had documented delivery dates.

The Applicant also conducted a logistic regression analysis for the primary efficacy endpoint; this analysis adjusted for additional covariates of gestational age at first dose, maternal age, cervical length, BMI and race.

Table 7 Study 302: Incidence of Preterm Birth (ITT Population, Applicant's Analysis)

Gestational Age at Delivery	Placebo	Progesterone Gel	Progesterone Gel vs. Placebo			
	N = 224	N = 235	Diff.	Relative Risk [a]	p-value [a] (CMH)	p-value [b] (Logistic)
	n (%)	n (%)	%	RR (95% CI)		
≤ 27 6/7 weeks	21 (9.4%)	12 (5.1%)	-4.3%	0.55 (0.28, 1.08)	0.075	0.133
≤ 32 6/7 weeks [c]	**34 (15.2%)**	**21 (8.9%)**	**-6.3%**	**0.56 (0.33, 0.93)**	**0.022**	**0.044**
≤ 34 6/7 weeks	50 (22.3%)	34 (14.5%)	-7.8%	0.61 (0.41, 0.90)	0.012	0.021
≤ 36 6/7 weeks	74 (33.0%)	71 (30.2%)	-2.8%	0.89 (0.68, 1.15)	0.377	0.532
≥ 37 weeks	150 (67.0%)	164 (69.8%)	2.8%	1.06 (0.93, 1.19)	0.377	0.532

[a] Cochran-Mantel-Haenszel test adjusted for primary pooled study site and risk strata
[b] Logistic regression with primary pooled study site, risk strata, treatment group, and covariables consisting of gestational age at first dose (weeks), maternal age (yrs), cervical length (cm), BMI (kg/m²), and race
[c] **Primary Efficacy Endpoint**
Source: Adapted from Applicant's Study Report for Study 302, Tables 11-7 and 11-9

Division Comments

- **The Applicant's primary analysis adjusted only for pooled study site and risk strata and found statistically significant differences favoring progesterone gel in the incidence of preterm birth at ≤ 32 6/7 weeks and ≤ 34 6/7 weeks. Differences at the other gestational ages evaluated were not statistically significant, but the trend favored progesterone gel in all cases. Results (also shown in Table 7) according to the logistic model were qualitatively similar, with statistically significant results favoring progesterone gel for delivery at ≤ 32 6/7 weeks and 34 6/7 weeks, and point estimates that favored progesterone gel at all weeks evaluated.**
- **The FDA statistical reviewer noted that the stratified CMH test used by the Applicant may not be appropriate when there are no or very few events across strata, as this makes it difficult to obtain a reliable point estimate. In fact, there were some sites that had no preterm deliveries, a fact that resulted in wide variations in efficacy, with results from some sites strongly favoring progesterone gel and others strongly favoring placebo. This suggests a treatment by site interaction. Because there were neither sufficient sample size in each strata, nor a consistent treatment effect among strata, use of the CMH test to adjust for pooled study site and risk strata does not appear to be appropriate.**
- **In addition, the background frequency of preterm birth was not comparable over US and non-US regions (discussed further below Table 8).**

The FDA's analysis is shown in Table 8. This analysis is based on a CMH test that adjusted only for region (US vs. non-US). A logistic regression model was also used that included terms for region, risk strata and treatment group, with additional covariates of gestational age at first dose, maternal age, cervical length, BMI and race. So the difference between the Applicant's and FDA's analyses lies in how the contributions of different sites are assessed, with the Applicant pooling study sites, and FDA evaluating US vs. non-US contributions.

Table 8 Study 302: Incidence of Preterm Birth (ITT Population, FDA Analysis)

Gestational Age at Delivery	Placebo	Progesterone Gel	Progesterone Gel vs. Placebo			
	N = 224	N = 235	Diff.	Relative Risk [a]	p-value [b] (CMH)	p-value [c] (Logistic)
	n (%)	n (%)	%	RR (95% CI)		
≤ 27 6/7 weeks	21 (9.4%)	12 (5.1%)	-4.3%	0.56 (0.29, 1.10)	0.071	0.157
≤ 32 6/7 weeks	**34 (15.2%)**	**21 (8.9%)**	**-6.3%**	**0.62 (0.37, 1.02)**	**0.033**	**0.056**
≤ 34 6/7 weeks	50 (22.3%)	34 (14.5%)	-7.8%	0.68 (0.47, 0.99)	0.023	0.030
≤ 36 6/7 weeks	74 (33.0%)	71 (30.2%)	-2.8%	0.95 (0.73, 1.22)	0.472	0.585
≥ 37 weeks	150 (67%)	164 (69.8%)	2.8%	1.06 (0.95, 1.19)	0.472	0.585

Bold = Primary Efficacy Endpoint
[a] Adjusted for region, cervical length, and maternal age
[b] Adjusted for region and risk strata
[c] Logistic regression adjusted for region, risk strata, treatment, gestational age at first dose (weeks), maternal age, cervical length, BMI (kg/m²), and race
Source: FDA Statistical Analysis and Review

Division Comment

Results based on the CMH test adjusted by region (US versus non-US) continue to show significance (p=0.033) at ≤ 32 6/7 week, and ≤ 34 6/7 weeks. After adjusting for additional covariables of maternal age and cervical length, progesterone gel was not associated with significant reduction in the preterm birth, as shown by the CI of the point estimate of the relative risk (RR) that includes one. This is also supported by the non-significant treatment effect (p = 0.056) based on the logistic regression analysis adjusting for additional variables at ≤ 32 6/7 weeks. Note that, to support drug approval based on a single study, efficacy evidence must be highly statistically persuasive.

The FDA statistical reviewer computed preterm birth rates by treatment arm and the Applicant's primary pooled site grouping. These results, shown in Table 9, demonstrate great variation in the rate of preterm birth (as evidenced by the rates in the placebo groups at different locations) and great heterogeneity in the treatment effect seen for progesterone gel.

Table 9 Study 302: Rate of Preterm Birth by Treatment and Primary Pooled Site (ITT Population, Applicant's Analysis)

Primary Pool Site	Placebo			Progesterone Gel			Diff.**
	n*	N#	%	n*	N#	%	
United States							
Midsouth/Midwest/West	5	22	22.7%	3	28	10.7%	-12.0%
Detroit MI	5	24	20.8%	5	25	20.0%	-0.8%
Mideast	4	26	15.4%	6	27	22.2%	6.8%
New York/New Jersey	5	27	18.5%	4	27	14.8%	-3.7%
India							
Andhra Pradesh	0	22	0.0%	2	23	8.7%	8.7%
Tamil Nadu/Gujarat	2	26	7.7%	0	24	0.0%	-7.7%
Maharashtra	4	28	14.3%	0	26	0.0%	-14.3%
Rest of the World							
Non Ukraine/India	8	23	34.8%	1	27	3.7%	-31.1%
Ukraine	1	26	3.8%	0	28	0.0%	-3.8%
Total	34	224	15.2%	21	235	8.9%	-6.2%

* n and % are the number and % of births ≤ 32 6/7 weeks gestation, N is total number of subjects
** Diff. is the difference represented by progesterone gel % minus placebo %.
Source: Applicant's Study Report for Study 302, Table 11-12

Division Comments

- **While the background rate of preterm birth at ≤ 32 6/7 weeks of gestation is fairly comparable across US sites (15-23%), far greater variation is seen in non-US sites, which range from no preterm births in placebo subjects at three Indian sites, to 35% in pooled non-US sites that exclude India and the Ukraine.**
- **The treatment effect also shows marked heterogeneity, even within the US, where the difference in the preterm birth rate in placebo-treated subjects compared to those treated with progesterone gel ranged from -12% (favoring progesterone gel) to +7% (favoring placebo). Similarly, three large Indian sites showed pooled results that favored placebo by 9%, while other non-US sites showed treatment effects favoring progesterone gel that ranged from -4% to -31%.**
- **These results suggest that FDA's analysis, pooling data by US vs. non-US sites, rather than by the Applicant's primary pooled site grouping may be more descriptive of efficacy, particularly efficacy within the US, which is the major concern.**

The logistic regression analysis done by FDA indicated that there were significant covariate terms in the model (as evidenced by a p-value < 0.10), including an interaction between treatment and region:

- Cervical length (p = 0.0001)
- Maternal age (p = 0.013)
- BMI (p = 0.019)
- Treatment (p = 0.02)
- Treatment by region interaction (p = 0.039)
- Risk strata (p = 0.062)

These results led the FDA statistical reviewer to prepare a stratified analysis of efficacy, shown in Table 10. These results fail to provide statistically significant evidence of efficacy

at any gestational age evaluated in the US population, although the trend favored progesterone gel. The risk ratios were close to unity, indicating no association of progesterone gel with a reduction of preterm births. However, results continue to show statistically significant evidence in support of progesterone gel for deliveries at ≤ 32 6/7 and ≤ 34 6/7 weeks in the non-US region.

Table 10 Study 302: Incidence of Preterm Birth Stratified by Region (ITT Population, FDA Analysis)

Gestational Age at Delivery	Placebo	Progesterone Gel	Progesterone Gel vs. Placebo			
			Diff.	Relative Risk [b]	p-value (CMH)	p-value[c] (Logistic)
	n (%)	n (%)	%	RR (95% CI)		
US Sites	**N = 99**	**N = 107**				
≤ 27 6/7 weeks	12 (12.1%)	9 (8.4%)	-3.7%	0.69 (0.31, 1.55)	0.380	0.627
≤ 32 6/7 weeks	**19 (19.2%)**	**18 (16.8%)**	**-2.4%**	**0.89 (0.50, 1.58)**	**0.659**	**0.688**
≤ 34 6/7 weeks	29 (29.3%)	27 (25.2%)	-4.3%	0.87 (0.57, 1.33)	0.514	0.439
≤ 36 6/7 weeks	41 (41.4%)	45 (42.1%)	0.7%	1.03 (0.76, 1.40)	0.926	0.968
Non-US Sites	**N = 125**	**N = 128**				
≤ 27 6/7 weeks	9 (7.2%)	3 (2.3%)	-4.9%	0.30 (0.08, 1.11)	0.070	0.154
≤ 32 6/7 weeks	**15 (12.0%)**	**3 (2.3%)**	**-9.7%**	**0.20 (0.06, 0.67)**	**0.003**	**0.016**
≤ 34 6/7 weeks	21 (16.8%)	7 (5.5%)	-11.3%	0.34 (0.15, 0.78)	0.004	0.017
≤ 36 6/7 weeks	33 (26.4%)	26 (20.3%)	-6.1%	0.74 (0.47, 1.16)	0.253	0.395

Bold = Primary Efficacy Endpoint
[b] Adjusted for Cervical Length and Maternal Age
[c] Logistic regression with risk strata, treatment, gestational age at first dose (weeks), maternal age (yrs), cervical length (cm), BMI (kg/m²), and race
Source: FDA Statistical Analysis and Review

Division Comment

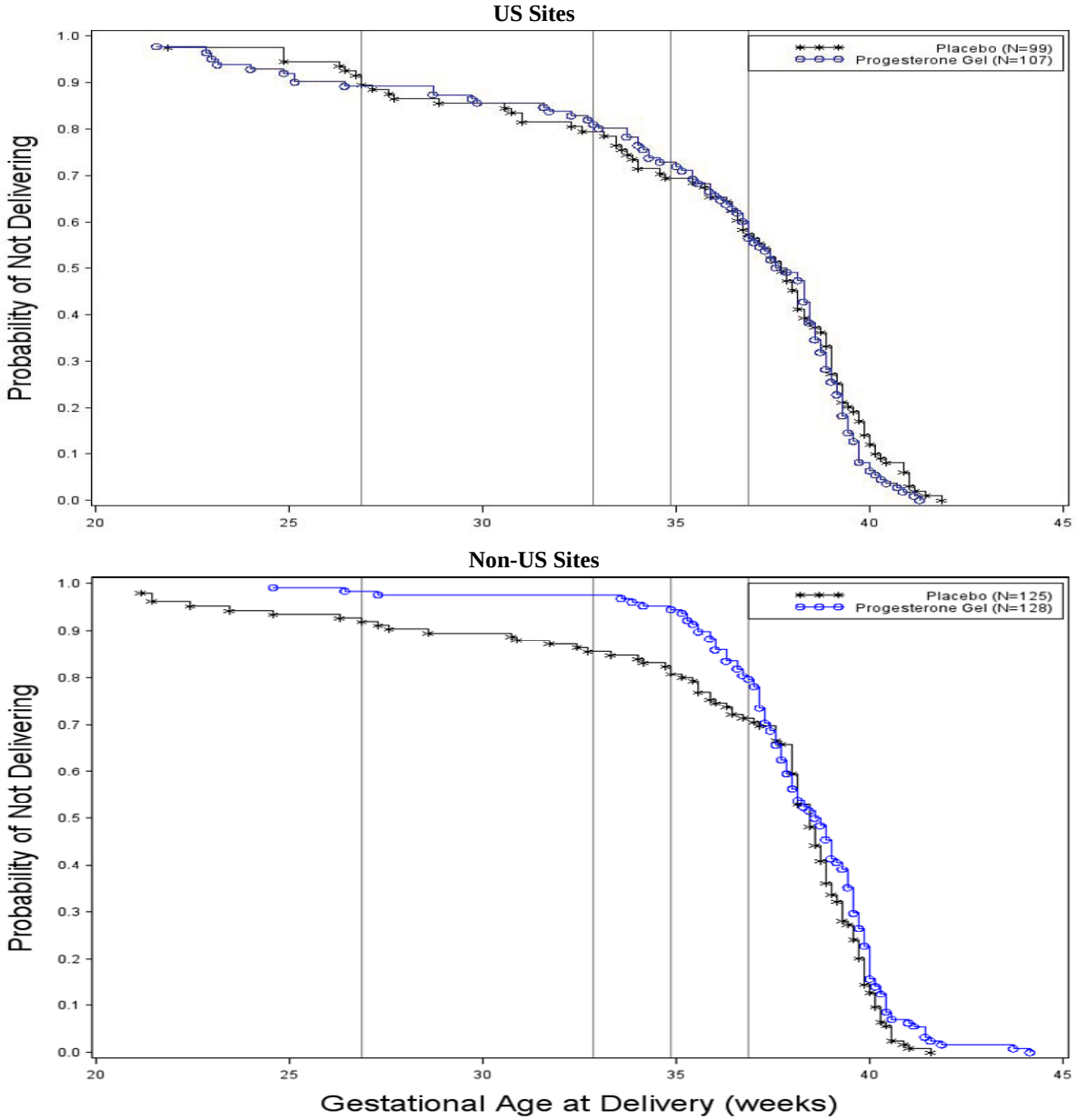
The Applicant initially attributed the lack of efficacy in the US population to higher rates of noncompliance in the US. However, the FDA statistical reviewer's analysis using the mITT population, which excluded subjects with < 80% compliance, produced similar results, with no statistically significant results in the US population at any of the gestational ages evaluated. Results in the non-US population continued to show statistical significance in favor of progesterone gel at gestational ages of ≤ 32 6/7 weeks and ≤ 34 6/7 weeks.

As requested by FDA, the Applicant prepared Kaplan-Meier curves of time to delivery for gestational age at delivery by treatment group, and for time from first dose by treatment group. In the former, there was some separation of the curves between 28 to 37 weeks of gestation, indicating that more women who received progesterone gel remained pregnant until term, the overall log-rank significance test was not significant (p = 0.42). In the latter Kaplan-Meier analysis, there was a similar non-significant trend for a higher probability of remaining pregnant for women treated with progesterone gel during the interval between about five to 16 weeks on treatment. However, after 16 weeks on treatment, the curves crossed. Depending on gestational age at entry, this would correspond to about 36 to 39 weeks of gestation, or essentially term.

In addition, because of the range of gestational ages at which subjects could initiate treatment, Kaplan-Meier curves for gestational age at delivery were prepared by the FDA statistical reviewer to adjust for the left truncation introduced by this staggered entry. The

result of this adjusted analysis was similar to that in the Applicant's initial analysis. The FDA reviewer also prepared separate Kaplan-Meier curves, accounting for staggered study entry, for the US and non-US regions (Figure 1).

Figure 1 Study 302: Kaplan-Meier Curve for Gestational Age at Delivery, Adjusting for Entry Time, by Region (ITT Population, FDA Analysis)



Note: Log-rank test for US Sites p-value=0.3667, for non-US Sites p-value=0.1012
Note: The vertical lines represent the preterm delivery cut-off points for 27 6/7, 32 6/7, 34 6/7, and 36 6/7 weeks
Source: FDA Statistical Analysis and Review

Division Comments

- **The Kaplan-Meier curves for US subjects show minimal difference between the progesterone gel and placebo arms at all time points. The curves for non-US subjects show separation at all time points between 28 and 37 weeks of gestation. The p-value for the log rank test was not significant in either region, however.**
- **It can be seen that the first preterm deliveries occurred much earlier in the placebo subjects in the non-US region (by Week 21), while the first preterm delivery in the progesterone gel arm did not occur until Week 25. In addition, there were no preterm deliveries in the progesterone gel arm in the non-US region between Weeks 27 and 33; this is a very different pattern from that observed in the US, where both treatment arms had preterm deliveries occurring throughout gestation.**
- **Further FDA analyses by region accounting for time on treatment were similar; there were no treatment differences between progesterone gel and placebo arms in the US, and in the non-US region, there were treatment differences favoring progesterone gel prior to Week 16 of treatment.**

Finally, the FDA statistical reviewer did a subgroup analysis of efficacy by country as a sensitivity analysis, shown in Table 11. As previously suggested in Table 9, there are marked differences among countries in preterm birth rates at various gestational ages and in the effects of treatment with progesterone gel.

Table 11 Study 302: Gestational Age at Delivery by Country (ITT Population, FDA Analysis)

Gestational Age at Delivery by Country	Placebo			Progesterone Gel			Diff. **
	n*	N#	%	n*	N#	%	
So. Africa							
≤ 32 6/7 weeks	**4**	**11**	**36.4%**	**0**	**10**	**0.0%**	**-36.4%**
≤ 34 6/7 weeks	5	11	45.5%	2	10	20.0%	-25.5%
≤ 36 6/7 weeks	5	11	45.5%	3	10	30.0%	-15.5%
Belarus							
≤ 32 6/7 weeks	**3**	**6**	**50.0%**	**0**	**5**	**0.0%**	**-50.0%**
≤ 34 6/7 weeks	3	6	50.0%	0	5	0.0%	-50.0%
≤ 36 6/7 weeks	3	6	50.0%	0	5	0.0%	-50.0%
India							
≤ 32 6/7 weeks	**6**	**76**	**7.9%**	**2**	**73**	**2.7%**	**-5.2%**
≤ 34 6/7 weeks	10	76	13.2%	4	73	5.5%	-7.7%
≤ 36 6/7 weeks	19	76	25.0%	18	73	24.7%	-0.3%
Ukraine							
≤ 32 6/7 weeks	**1**	**26**	**3.8%**	**0**	**28**	**0.0%**	**-3.8%**
≤ 34 6/7 weeks	2	26	7.7%	0	28	0.0%	-7.7%
≤ 36 6/7 weeks	4	26	15.4%	2	28	7.1%	-8.2%
United States							
≤ 32 6/7 weeks	**19**	**99**	**19.2%**	**18**	**107**	**16.8%**	**-2.4%**
≤ 34 6/7 weeks	29	99	29.3%	27	107	25.2%	-4.1%
≤ 36 6/7 weeks	41	99	41.4%	45	107	42.1%	0.6%

* n and % are the number and % of births at gestation age
** Diff. is the difference represented by progesterone gel % of preterm births minus placebo % of preterm births
Bold = Primary Efficacy Endpoint
Source: FDA Statistical Analysis and Review

Division Comments

- **South Africa (with 11 placebo subjects and 10 progesterone gel subjects) and Belarus (with six placebo subjects and five progesterone gel subjects) had very high rates of preterm birth in the placebo subjects, and no preterm births at ≤ 32 6/7 weeks among progesterone gel-treated subjects. This resulted in a large treatment effect favoring progesterone gel in these countries.**

- **In contrast, preterm delivery rates among placebo subjects in India and Ukraine are very low, but still 4-8% higher than those among progesterone gel-treated subjects.**

- **In the US, preterm delivery rates in placebo-treated subjects are generally similar to those seen for placebo subjects in other clinical trials studying preterm birth rates. This suggests that these rates are likely to be fairly representative of the US background rate for preterm birth among women at high risk of preterm birth. The treatment effect seen in the US was small, ranging from 0.6% (favoring placebo) at < 37 weeks, to -4% and < 35 weeks, and -2% at the primary endpoint of < 33 weeks.**

- **When the FDA statistical reviewer did a further sensitivity analysis excluding South Africa and Belarus, results were no longer significant at any gestational age evaluated in the non-US population, or in the overall ITT population (US and non-US). Because these two sites were relatively small compared to the total subject population (contributing less than 8% of subjects in each arm), it is unlikely that this is related to lack of statistical power. Rather, it appears that the overall efficacy results may be driven by these discrepant sites.**

4.1.3.2 Main Secondary Efficacy Analysis

Mortality/Morbidity Composite Score

The most important secondary endpoint was the neonatal mortality/morbidity composite score. The Applicant analyzed this in several ways as described in Section 3.1.4.2, and overall results are shown in Table 12. Although there was a trend toward lower rates of mortality/morbidity among progesterone-gel treated subjects, none of the analyses were statistically significant.

Table 12 Study 302: Mortality/Morbidity Composite Score (ITT Population)

Analysis	Value	Placebo n (%)	Procheive n (%)	p-value
Any Morbidity or Mortality Event	Yes	28 (12.5%)	18 (7.7%)	0.088*
0-4 Point Scale Composite Score¶	0	192 (86%)	217 (92%)	0.133*
	1	11 (5%)	5 (2%)	
	2	7 (3%)	2 (<1%)	
	3	0 (0%)	3 (1%)	
	4	10 (4%)	8 (3%)	
	Mean	0.3	0.2	0.288**
0-4 Point Scale with NICU Days Composite Score#	0	168 (75%)	194 (83%)	0.103*
	1	11 (5%)	6 (3%)	
	2	17 (8%)	19 (8%)	
	3	14 (6%)	8 (3%)	
	4	10 (4%)	8 (3%)	
	Mean	0.6	0.4	0.102**
0-6 Point Scale Composite Score§	0	192 (86%)	217 (92%)	0.113*
	1	11 (5%)	2 (2%)	
	2	7 (3%)	2 (<1%)	
	3	0 (0%)	0 (0%)	
	4	0 (0%)	3 (1%)	
	5	0 (0%)	0 (0%)	
	6	10 (4%)	8 (3%)	
	Mean	0.4	0.3	0.404**

* Cochran-Mantel-Haenszel (CMH) and **Analysis of variance (ANOVA); CMH and ANOVA both have terms for primary pooled study site, risk strata, and treatment group.

¶ 0=no events; 1=one event for (RDS, grade III or IV IVH, proven sepsis, PVL, NEC, BPD) and no perinatal mortality; 2=two events and no perinatal mortality; 3=three or more events and no perinatal mortality; 4=perinatal mortality.

0=no events; 1=one event for (RDS, grade III or IV IVH, proven sepsis, PVL, NEC, BPD) or <5 days in the NICU and no perinatal mortality; 2=two events or between 5 and 20 days in the NICU and no perinatal mortality; 3=three or more events or >20 days in the NICU and no perinatal mortality; 4=perinatal mortality.

§ 0=no events; 1=one event for (RDS, grade III or IV IVH, proven sepsis, NEC, BPD) and no perinatal mortality; 2=two events and no perinatal mortality; 3=three events and no perinatal mortality; 4=four events and no perinatal mortality; 5=five events and no perinatal mortality; 6=Perinatal mortality.

Source: Applicant's Study Report for Study 302, Table 11-15

Division Comment

The 0-4 point composite score without NICU days was the specified primary analysis for this key secondary endpoint, but the Division had requested the 0-6 score be included as a sensitivity analysis.

4.2 Study 300

The ITT population included 611 subjects, 309 who received progesterone gel and 302 who received placebo. The SHCX group enrolled only nine subjects (four treated with progesterone gel, five with placebo).

In general, the two treatment groups were similar with regard to demographics.

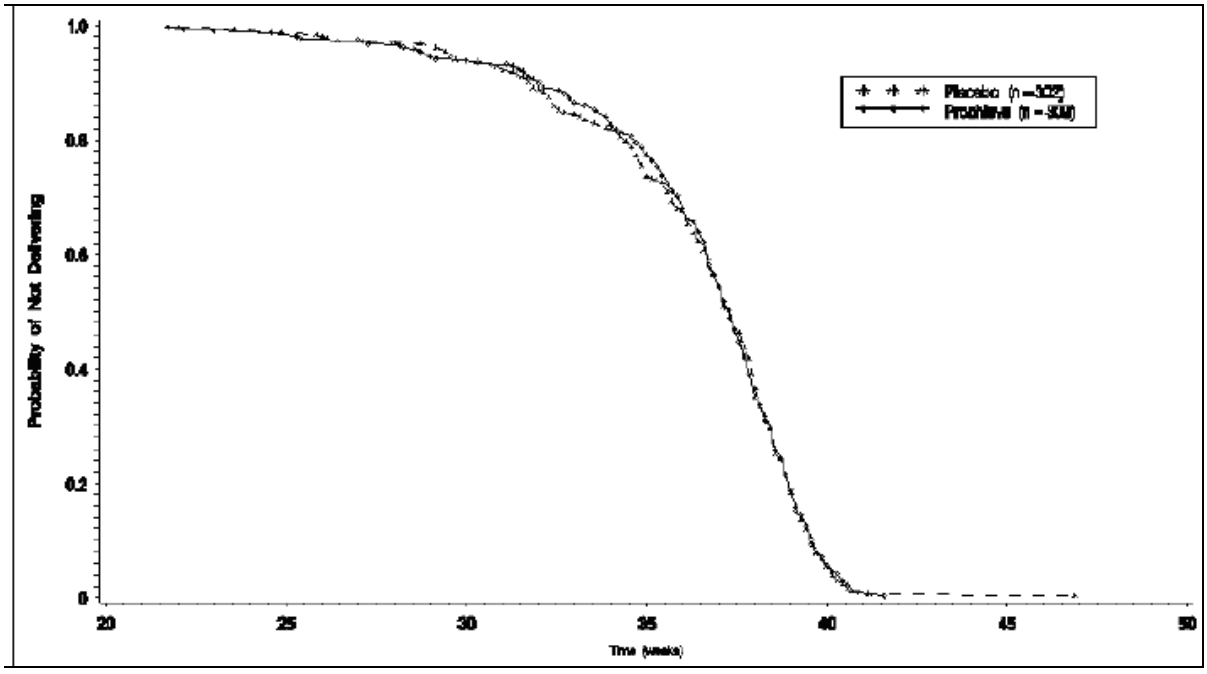
4.2.1 Efficacy Findings

Efficacy findings for the ITT population are presented briefly; the results did not provide evidence of efficacy for progesterone gel in a population of women at risk due to a prior preterm birth. Additional, post hoc, efficacy findings more relevant to the population at risk due to a short cervix are presented.

4.2.1.1 Primary Efficacy Endpoint and Analysis

The proportion of women with a prior preterm birth who had a delivery before 32 0/7 weeks gestation did not differ between treatment groups (10% of progesterone gel subjects and 11% of placebo subjects, p = 0.694. The Kaplan Meier curve for probability of delivery by week of gestation shows no difference by treatment group (Figure 2).

Figure 2 Study 300: Probability of Delivery by Gestational Age and Treatment (ITT)



Source: Applicant's Study Report for Study 300, Figure 14.2.1

Division Comments

- **The trial failed on the primary efficacy endpoint.**
- **The Applicant reported an "unexpected finding" that subjects in the progesterone gel group had less cervical shortening between baseline and Week 28 than did placebo subjects. For this reason, the Applicant conducted a quartile analysis of the primary endpoint.**

4.2.1.2 Post Hoc Efficacy Analyses Results

Results stratified by quartile of baseline cervical length were also non-significant in all four quartiles; however, the Applicant noted a 50% reduction in the frequency of preterm birth for progesterone gel subjects compared to placebo subjects in the quartile with the shortest cervical length (<3.2 cm, N = 172). These results are shown in Table 13. The Applicant further evaluated the overall risk of preterm birth in a subgroup of women with baseline cervical length ≤ 3 cm (N = 58 in each treatment arm); these results were not statistically significant and are displayed in Figure 3. Further post hoc analyses of a subgroup with cervical length ≤ 2.8 cm (N = 28 in progesterone gel arm, 30 in placebo arm) found a statistically significant reduction in deliveries at ≤ 32 0/7 weeks of gestation (p=0.002) and a statistically significantly lower morbidity/mortality composite score. Most of these women (aside from the nine in the SHCX group) also had a previous preterm birth.

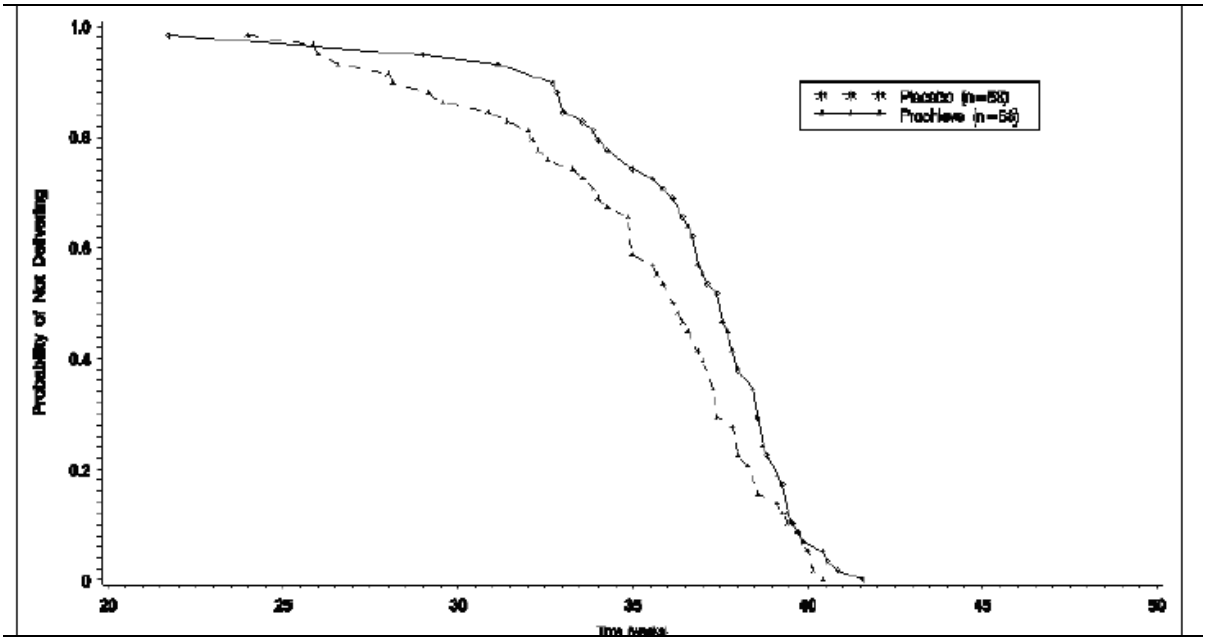
Table 13 Study 300: Gestational Age at Delivery by Quartiles of Cervical Length and Treatment (ITT Population)

	Quartile 1 (<3.2 cm)	Quartile 2 (≥3.2 to <3.6 cm)	Quartile 3 (≥3.6 to <4.2 cm)	Quartile 4 (≥4.2 cm)
Placebo (n/N [%])	12/89 (13.5%)	8/74 (10.8%)	5/69 (7.2%)	10/68 (14.7%)
Prochieve (n/N [%])	6/83 (7.2%)	10/72 (13.9%)	7/83 (8.4%)	7/71 (9.9%)
p-value[1]	0.218	0.622	1.000	0.444

[1] P-value for the overall test of raw proportions based on Fisher's exact test.

Source: Applicant's Study Report for Study 300, Table 11.4

Figure 3 Study 300: Probability of Delivery by Gestational Age and Treatment (Baseline Cervical Length ≤ 3 cm)



Source: Applicant's Study Report for Study 300, Figure 14.2.2.2x

Division Comment

The FDA reviewer also analyzed gestational age at delivery by baseline cervical length ≤ 2.5 cm, a common clinical definition of short cervix. There were 12 subjects in the progesterone

gel arm and 21 in the placebo arm in this subgroup. The proportion of preterm deliveries favored progesterone gel at the earlier gestational ages (≤ 27 6/7 weeks and ≤ 32 6/7 weeks) and favored placebo that later gestational ages (≤ 34 6/7 weeks and ≤ 36 6/7 weeks); none of these results reached statistical significance.

4.3 Additional Efficacy Analyses

In Study 302, the Applicant conducted an additional subgroup analysis stratified by cervical length (≤1.7 cm or >1.7 cm) as a sensitivity analysis. Additional post hoc subgroup analyses by region (US vs. non-US) were provided. After preliminary review of the study reports by the Division, additional subgroup analyses including subgroup analyses by race, age group, BMI, and region were requested by the Division. .In addition, the FDA reviewer did subgroup analyses stratified by race, BMI, maternal age and cervical length (≤1.8 cm or >1.8 cm). These were variables identified in the logistic regression model has having a statistically significantly impact on preterm delivery rates.

Division Comment

The Applicant chose the 1.7 cm cutpoint for cervical length as the median value, stating that 42% of the population was included in the group with cervical length ≤ 1.7 cm. However, 56% of the population was included when the cutpoint was set at 1.8 cm; this value was selected by the FDA statistical reviewer as a better representation of the median value and used as the cutpoint in the FDA analysis.

4.3.1 Efficacy by Race

Results stratified by race (Asian, Black, and Caucasian) are presented in Table 14. Kaplan-Meier curves adjusting for gestational age at entry were also calculated; the log rank tests for each curve (each representing a different racial group) were all non-significant.

Table 14 Study 302: Gestational Age at Delivery by Race (ITT Population, FDA Analysis)

Gestational Age at Delivery by Race	Placebo			Progesterone Gel			Difference **
	n*	N#	%	n*	N#	%	
Asian							
≤ 32 6/7 weeks	**6**	**77**	**7.8%**	**3**	**76**	**3.9%**	**-3.8%**
≤ 34 6/7 weeks	10	77	13.0%	5	76	6.6%	-6.4%
≤ 36 6/7 weeks	19	77	24.7%	20	76	26.3%	1.6%
Black							
≤ 32 6/7 weeks	**15**	**67**	**22.4%**	**12**	**76**	**15.8%**	**-6.6%**
≤ 34 6/7 weeks	22	67	32.8%	19	76	25.0%	-7.8%
≤ 36 6/7 weeks	31	67	46.3%	28	76	36.8%	-9.4%
Caucasian							
≤ 32 6/7 weeks	**10**	**69**	**14.5%**	**4**	**73**	**5.5%**	**-9.0%**
≤ 34 6/7 weeks	14	69	20.3%	8	73	11.0%	-9.3%
≤ 36 6/7 weeks	20	69	29.0%	20	73	27.4%	-1.6%

* n and % are the number and % of births at gestation age, N is total number of subjects
** Difference is progesterone gel % minus placebo %.
Source: FDA Statistical Analysis and Review

Division Comments

- **The preterm birth rate in placebo-treated subjects varies considerably by race. This may also reflect regional differences, as race was differentially distributed by region.**

- **Overall, there was minimal efficacy in Asian subjects, possibly because the background rate of prematurity was quite low.**
- **Efficacy in Caucasians and Blacks favored progesterone gel. In Caucasians, the treatment benefit was greatest in the earlier gestational ages, while in Blacks, efficacy increased as gestational age advanced.**

4.3.2 Efficacy by Maternal Age

Subjects were initially grouped by quartiles of maternal age. Based on similar preterm delivery rates in women between 22 to 30 years, subjects were then classified into three groups (< 22 years, 22 to <30 years, and ≥ 30 years old). Results are shown in Table 15. Kaplan-Meier curves were computed for each of the three age groups, the log rank test was statistically significant in favor of progesterone gel for women aged 22 to 30 years.

Table 15 Study 302: Gestational Age at Delivery by Age (ITT Population, FDA Analysis)

Gestational Age at Delivery by Age Group	Placebo			Progesterone Gel			Difference **
	n*	N	%	n*	N	%	
< 22 years							
≤ 32 6/7 weeks	**4**	**59**	**6.8%**	**10**	**62**	**16.1%**	**9.3%**
≤ 34 6/7 weeks	10	59	16.9%	16	62	25.8%	8.9%
≤ 36 6/7 weeks	18	59	30.5%	22	62	35.5%	5.0%
22 - < 30 years							
≤ 32 6/7 weeks	**22**	**115**	**19.1%**	**7**	**113**	**6.2%**	**-12.9%**
≤ 34 6/7 weeks	30	115	26.1%	11	113	9.7%	-16.4%
≤ 36 6/7 weeks	41	115	35.7%	22	113	19.5%	-16.2%
≥ 30 years							
≤ 32 6/7 weeks	**8**	**50**	**16.0%**	**4**	**60**	**6.7%**	**-9.3%**
≤ 34 6/7 weeks	10	50	20.0%	7	60	11.7%	-8.3%
≤ 36 6/7 weeks	15	50	30.0%	23	60	38.3%	8.3%

* n and % are the number and % of births at gestation age, N is total number of subjects
** Difference is progesterone gel % minus placebo %.
Source: FDA Statistical Analysis and Review

Division Comments

- **In young women (< 22 years old), the rate of preterm delivery at all gestational ages examined was higher in the progesterone gel arm.**
- **The treatment effect favoring progesterone gel was largest in women aged 22 to 30 years.**
- **It is unclear whether these discrepancies reflect chance variation or a difference by age in the underlying etiology of preterm birth, which is reflected in the differential impact of progesterone gel on the preterm birth rate.**

4.3.3 Efficacy by Body Mass Index

The effect of BMI was also assessed; subjects were categorized by BMI < 20 kg/m^2, 20 to < 30 kg/m^2 or ≥ 30 kg/m^2. Results are presented in Table 16. Log rank tests on the Kaplan-Meier curves computed for each of the three BMI categories were all non-significant.

Table 16 Study 302: Gestational Age at Delivery by BMI (ITT Population, FDA Analysis)

Gestational Age at Delivery by BMI	Placebo			Progesterone Gel			Difference **
	n*	N	%	n*	N	%	
< 20							
≤ 32 6/7 weeks	**5**	**46**	**10.9%**	**3**	**51**	**5.9%**	**-5.0%**
≤ 34 6/7 weeks	8	46	17.4%	5	51	9.8%	-7.6%
≤ 36 6/7 weeks	13	46	28.3%	19	51	37.3%	9.0%
20 - < 30							
≤ 32 6/7 weeks	**15**	**125**	**12.0%**	**10**	**121**	**8.3%**	**-3.7%**
≤ 34 6/7 weeks	24	125	19.2%	14	121	11.6%	-7.6%
≤ 36 6/7 weeks	40	125	32.0%	27	121	22.3%	-9.7%
≥ 30							
≤ 32 6/7 weeks	**13**	**48**	**27.1%**	**7**	**56**	**12.5%**	**-14.6%**
≤ 34 6/7 weeks	17	48	35.4%	13	56	23.2%	-12.2%
≤ 36 6/7 weeks	20	48	41.7%	23	56	41.1%	-0.6%

* n and % are the number and % of births at gestation age, N is total number of subjects
** Difference is progesterone gel % minus placebo %.
Source: FDA Statistical Analysis and Review

Division Comments

- **The treatment effect is greatest in heavier women (BMI > 30) for gestational ages < 35 weeks.**
- **In women with BMI < 20, the treatment effect crossed over between 35 and 37 weeks, favoring progesterone gel at earlier gestational ages, and favoring placebo for the proportion of preterm deliveries before 37 weeks.**

4.3.4 Efficacy by Cervical Length

The Applicant did a subgroup analysis by cervical length ≤1.7 cm or >1.7 cm. Results are displayed in Table 17.

Table 17 Study 302: Gestational Age at Delivery by Cervical Length, Cut at 1.7 cm (ITT Population, Applicant's Analysis)

	Gestational Age (Weeks)	Descriptive Statistic	Placebo	Prochieve	% Reduction Relative to Placebo**
Cervical Length ≤ 1.7 cm	≤27 6/7	N n (%) p-value*	94 14 (14.9%)	99 8 (8.1%) 0.119	-46%
	≤32 6/7	N n (%) p-value*	94 23 (24.5%)	99 15 (15.2%) 0.068	-43%
	≤34 6/7	N n (%) p-value*	94 31 (33.0%)	99 23 (23.2%) 0.075	-34%
Cervical Length > 1.7 cm	≤27 6/7	N n (%) p-value*	130 7 (5.4%)	136 4 (2.9%) 0.447	-37%
	≤32 6/7	N n (%) p-value*	130 11 (8.5%)	136 6 (4.4%) 0.182	-47%
	≤34 6/7	N n (%) p-value*	130 19 (14.6%)	136 11 (8.1%) 0.109	-43%

Source: Applicant's Study Report for Study 302, Table 14.2.1.5

The FDA statistical reviewer did subgroup analyses stratifying by cervical length ≤1.8 cm or >1.8 cm, which was the actual population median cervical length in Study 302.

Table 18 Study 302: Gestational Age at Delivery by Cervical Length, Cutpoint at 1.8 cm (ITT Population, FDA Analysis)

Gestational Age at Delivery by Cervical Length	Placebo		Progesterone Gel		Progesterone Gel vs. Placebo	
	n	%	n	%	Difference	p-value [a]
≤ 1.8 cm	**N = 130**		**N = 131**			
≤ 27 6/7 weeks	17	13.1%	9	6.9%	-6.2%	0.086
≤ 32 6/7 weeks	**28**	**21.5%**	**16**	**12.2%**	**-9.3%**	**0.037**
≤ 34 6/7 weeks	40	30.8%	25	19.1%	-11.7%	0.022
≤ 36 6/7 weeks	53	40.8%	49	37.4%	-3.4%	0.524
> 1.8 cm	**N = 94**		**N = 104**			
≤ 27 6/7 weeks	4	4.3%	3	2.9%	-1.4%	0.599
≤ 32 6/7 weeks	**6**	**6.4%**	**5**	**4.8%**	**-1.6%**	**0.617**
≤ 34 6/7 weeks	10	10.6%	9	8.7%	-2.0%	0.620
≤ 36 6/7 weeks	21	22.3%	22	21.2%	-1.2%	0.824

[a] Cochran-Mantel-Haenszel test adjusted for region
Source: FDA Statistical Analysis and Review

Division Comments

- **In the Applicant's analysis, the treatment effect favored progesterone gel at all gestational ages examined in both cervical length strata. However, none of the results attained statistical significance. The treatment effect was fairly consistent over both strata. The**

differences in preterm birth rates (as opposed to percent reductions relative to placebo) ranged from -2.5% to -9.8%

- **In the FDA analysis, there was minimal effect of progesterone gel in women with cervical length > 1.8 cm. However, the treatment effect, which favored progesterone gel at all gestational ages examined in women with cervical length ≤ 1.8 cm, was statistically significant at ≤ 32 6/7 and ≤ 34 6/7 weeks of gestation.**

4.4 Applicant's Evaluation of Chorioamnionitis

In the NDA filing letter of June 28, 2011, the Division notified the Applicant that the discrepancy in efficacy results across US and foreign sites was a potential review issue. Early in the review cycle, the Division requested the Applicant to submit any information it found pertinent to explain this difference in efficacy results. On September 15, 2011, the Applicant submitted additional analyses and commentary on regional differences in efficacy, concluding that regional differences could represent inherent variability attributable to low event rates across multiple subgroups. The Applicant also noted an imbalance in chorioamnionitis, with more cases diagnosed in the US, and more cases observed among subjects who delivered preterm before 33 weeks of gestation.

Chorioamnionitis is clearly associated with preterm labor and preterm birth. When diagnosed prior to delivery, diagnosis is made on the basis of a constellation of clinical symptoms. However, many cases remain subclinical and are not diagnosed prior to delivery.

Chorioamnionitis is a fairly common histopathologic finding when placental tissues, fetal membranes, and the umbilical cord are sent for examination. This examination is most often requested following premature births or deliveries with poor infant outcomes. Overall, in a large population of preterm infants in which >90% of placentas were sent for histopathologic analysis, the overall incidence of histologic chorioamnionitis was 31%[3]. There was an inverse relationship between gestational age at delivery and chorioamnionitis, with rates ranging from 66% of liveborn infants delivered at 20-24 weeks to 16% of those delivered at 34 weeks.

The Applicant made an additional submission on November 4, 2011, which provided further evaluation of the impact of chorioamnionitis. In this submission, an analysis was done that excluded cases of clinical chorioamnionitis from the efficacy evaluation. In the overall ITT population, 20 cases were excluded, 17 of these in the US. Results remained statistically significant in the total population and the non-US region, and remained non-significant in the US region, although the treatment effect at < 33 weeks was slightly higher than that obtained without excluding these subjects (-8.1%, 95% CI -17.8 to 1.5%). The Applicant noted that the treatment effect at the primary endpoint of ≤ 32 6/7 weeks was similar in the US and non-US regions (-9.0%) when clinical chorioamnionitis was excluded, and suggested that regional differences may reflect, in part, disproportionate rates of chorioamnionitis in the US sites.

Division Comments

- **The diagnosis of chorioamnionitis, whether clinically or by histopathology, is not based upon standard and internationally accepted diagnostic criteria, so it is not possible to evaluate whether the threshold for diagnosis differed between the US and non-US regions.**

- **In addition, there is no accepted screening test for chorioamnionitis in mid-pregnancy, so it is not possible to identify most of these cases at or prior to 20 weeks gestation, when treatment with progesterone gel would begin.**

4.5 Overall Summary of Efficacy

The information and data in this application do not support the efficacy of progesterone gel compared with placebo in reducing the risk of preterm births before 33 completed weeks of gestation among women with a short cervical length. This conclusion is based on the analysis of data from a single multinational study (Study 302) conducted in ten countries.

Supplemental information from Study 300 did not provide adequate evidence to be considered supportive of efficacy, as the "short cervix" subgroups analyzed included only nine to 116 subjects in total, depending on the specific definition used for short cervix. Cervical length ≤ 2.5 cm is generally considered to represent a clinically meaningful short cervix and a risk factor for preterm birth; only 33 subjects in Study 300 fell into this range. Although the Applicant presented analyses based on pooled data from Studies 300 and 302, this was not felt to be statistically sound due to the differences in the study populations and initiation of treatment.

The FDA analyses accounted for quantitative heterogeneity (a region by treatment interaction) and additional covariates of gestational age at first dose, maternal age, cervical length, body mass index, and race. Although these demographic and baseline characteristics appeared to be balanced between treatment arms, they differed notably by region. Subgroup analyses that stratified for these covariates individually showed that the treatment effect differed over strata. After adjusting for these covariates, the FDA analysis in the overall population indicated that, although subjects treated with progesterone gel demonstrated a difference in the risk of preterm births at ≤ 32 6/7 weeks of -6.3% compared to the risk in placebo subjects, the result was not statistically significant. The Applicant's analysis (a CMH test stratified by primary pooled study site and risk strata) was found not to be appropriate due to insufficient sample size in each stratum and inconsistent treatment effects among strata.

Most importantly, the FDA analyses also indicated that progesterone gel was not associated with a reduction in preterm birth in the US subjects at any gestational age, with a treatment difference of only -2.4% in favor of progesterone gel compared to placebo at ≤ 32 6/7 weeks gestation. The CI around this point estimate of the risk difference indicated that the difference is not statistically significant. In addition, the key secondary endpoint of neonatal mortality and morbidity did not show a statistically significant treatment effect of progesterone gel.

The most important statistical analysis issue noted in this application regarded the regional heterogeneity in efficacy. The treatment effect varied substantially across countries and there was a significant region by treatment quantitative interaction. The FDA analysis addressed this issue by pooling US vs. non-US as a region and adjusting for the region in the CMH analysis and the logistic regression model. An analysis by region and additional sensitivity analysis were also conducted by the FDA. The results are summarized as follows:

- In the US population, which made up about 45% of the total study population, the treatment difference was -2.4% and was not statistically significant at the primary

efficacy endpoint gestational age of ≤ 32 6/7 weeks (p=0.68). Similar results were seen at gestational ages of ≤ 34 6/7 and ≤ 36 6/7 weeks.

- Sensitivity analysis excluding two countries with the highest treatment differences (only 7% of the subjects) resulted in a non-statistically significant overall treatment difference for progesterone gel at any time points.

From a statistical perspective, the evidence from this single study does not support the efficacy of progesterone gel 8% for the prevention of preterm deliveries among women with a short cervical length.

From a clinical perspective, it does not appear that the Applicant has identified a population of US women who are likely to benefit from the use of progesterone gel to reduce their risk of preterm birth. Even if some of the efficacy disparity seen between US and non-US sites is attributable to differential rates of chorioamnionitis, this is not a factor that can be addressed in labeling so as to target a specific population in whom the drug is likely to be of benefit. Subjects with clinical symptoms suggestive of chorioamnionitis were excluded from the trial, but, at present, there is no way to identify women likely to develop chorioamnionitis prior to the gestational age at which progesterone gel would be started.

5. Safety Finding from Progesterone Gel Clinical Trials

5.1 Overview of the Safety Database for Progesterone Gel

The progesterone gel safety database includes data from one phase 1 PK study, and two phase 3 studies that enrolled a total of 1,119 pregnant women. The phase 1 study enrolled 23 women and there were a total of 1,096 in the phase 3 studies. Although the initiation of treatment varied slightly by study, all continued treatment until the earlier of delivery of 36 6/7 weeks gestation.

5.2 Safety Findings from the Progesterone Gel Clinical Development Program

5.2.1 Deaths

No maternal deaths were reported in the clinical development program. Intrauterine fetal deaths (IUFD), stillbirths, neonatal and infant deaths were reported only in the two phase 3 trials. These terms are defined as follows:

- IUFD – fetal demise after 20 weeks gestation and prior to the onset of labor
- Stillbirth – fetal demise after 20 weeks gestation and during labor
- Neonatal death – of a liveborn, up through 28 days post-delivery
- Infant death – at > 28 days post-delivery

Table 19 Fetal, Neonatal and Infant Deaths by Study Drug

	Study 302				Study 300				Pooled Population			
	Progesterone N=224		Placebo N=235		Progesterone N=321		Placebo N=316		Progesterone N=545		Placebo N=551	
	n	%	n	%	n	%	n	%	n	%	n	%
IUFD	2	0.9	2	0.9	4	1.2	1	0.3	6	1.1	3	0.5
Stillbirth	3	1.3	4*	1.3	1	0.3	3*	0.6	4	0.7	5**	0.9
Neonatal Death	3	1.3	5	2.2	6*	1.6	7*	1.9	8**	1.5	11**	2.0
Infant Death	0	0	0	0	3	0.9	3	0.9	3	0.6	3	0.5

* one occurred prior to receiving study drug; percents exclude these cases
**total excludes cases occurring prior to study drug
Source: Based on Applicant's Summary of Clinical Safety and Table 6.2.1-1 Integrated Summary of Safety

Division Comments

- **Combining IUFD and stillbirths in the pooled population, the rate of fetal loss remains slightly greater in the progesterone gel arm (1.8% vs. 1.4%). It is unlikely that this represents a true difference in risk for fetal loss.**
- **For overall fetal and neonatal deaths, the rate was 3.9% in the pooled progesterone gel arm and 4.0% in the placebo arm. This does not reflect a clinically meaningful benefit of progesterone gel on overall mortality.**

5.2.2 Non-fatal Serious Adverse Events

Phase 1 studies

Fifteen serious adverse events (SAEs) in four subjects were reported in the phase 1 clinical trial; none was considered to be treatment-related by the investigators. These included: premature rupture of membranes (PROM), preterm labor and preterm delivery in one subject; motor vehicle accident, preterm labor and preterm delivery in one subject; cervical insufficiency, PROM, preterm labor and preterm delivery in one subject; and nausea, vomiting, headache, dehydration and preterm contractions in the fourth subject.

Phase 3 studies

In Study 302, SAEs occurred in 41% of progesterone gel subjects and 44% of placebo subjects. These rates are high because many common pregnancy complications, including preterm birth and preterm labor, are included in the definition of SAEs. SAEs that occurred in at least 1% of subjects and with greater frequency in the progesterone gel arm compared to the placebo arm are shown in Table 20

Table 20 Selected SAEs in Phase 3 Studies

Preferred Term	Study 300			
	Progesterone Gel N = 321		Placebo N =316	
	n	%	n	%
Uterine contractions abnormal	17	5	10	3
Cervical incompetence	8	3	6	2
Intrauterine death/stillbirth	5	2	4	1
Abdominal pain lower	4	1	0	0
	Study 302			
	Progesterone Gel N =224		Placebo N =235	
Cervix disorder	21	9	12	5
Premature rupture of membranes	12	5	9	4
Cervix cerclage procedure	9	4	5	2
Fetal distress syndrome	6	3	3	1
Cervical incompetence	4	2	3	1
Pregnancy-induced hypertension	4	2	1	< 1

Source: Applicant's Study Reports for Study 300 (Table 12.4) and Study 302 (Table 14.3.1.5**).**

Division Comment
The main SAE of interest is cervical incompetence, which occurred slightly more frequently in the progesterone gel subjects.

5.2.3 Discontinuations due to Adverse Events

No subjects in the phase 1 study discontinued from the study due to an adverse event (AE). In the phase 3 studies, discontinuations due to AEs were similar over treatment arms (18% each in the pooled dataset) but varied greatly by study: in Study 300, less than 2% of subjects in either arm discontinued due to an AE, but in Study 302, 27-29% did so.

5.2.4 Common Adverse Events

The most common maternal AEs for each study are reported in Table 21, based on AEs that occurred in at least 5% of subjects and were more common among subjects in the progesterone gel.

Table 21 Selected Common Adverse Events in the Phase 3 Studies (ITT Population)

Preferred Term	Study 300		Study 302	
	Progesterone (N=321) n (%)	Placebo N=316 n (%)	Progesterone (N=224) n (%)	Placebo (N=235) n (%)
At least one AE	261 (81)	263 (83)	159 (68)	155 (69)
Premature labor	63 (20)	66 (21)	17 (7)	33 (15)
Uterine contractions abnormal	46 (14)	34 (11)	34 (14)	40 (18)
PROM	22 (7)	19 (6)	13 (6)	10 (4)
Vulvovaginal mycotic + vaginal infection	25 (8)	32 (10)	32 (14)	22 (9)
Urinary tract infection	36 (11)	18 (6)	11 (5)	9 (4)
Abdominal pain + lower abdominal pain	47 (15)	37 (12)	13 (6)	11 (5)
Dyspepsia	24 (8)	26 (8)	15 (6)	12 (5)
Cervix disorder	20 (6)	35 (11)	23 (10)	17 (8)
Back pain	35 (11)	27 (9)	12 (5)	11 (5)
Headache	31 (10)	29 (9)	16 (7)	18 (8)
Uterine contractions during pregnancy	40 (13)	37 (12)		
Edema peripheral	22 (7)	15 (5)		
Gestational diabetes			14 (6)	9 (4)
Bacterial vaginitis			12 (5)	6 (3)

Source: Applicant's Summary of Clinical Safety, Adapted from Table 2.7.4-4

Division Comments

- **As shown in the table above, certain events appear to have been reported only in one of the studies; it is not clear whether this reflects a true difference in events, or differences in ascertainment of AEs.**
- **Overall, there do not appear to be major differences in the rates of common AEs between progesterone gel-treated and placebo-treated subjects.**

5.2.5 Two-Year Infant Follow-up

As requested by the Division, the Applicant assessed the offspring of treated women in Study 300 at 6-, 12- and 24-month follow-up visits. Evaluations included measurement of head circumference, weight and length, documentation of congenital abnormalities, chronic morbid conditions and results on the Denver Developmental Screening Test II (DDST II). Of 321 infants born to progesterone gel-exposed women and 316 born to placebo-exposed women, the proportion that was followed was similar in each exposure group at the 6-month visit, and slightly higher in the placebo group at subsequent visits. About two-thirds of all infants were evaluated at six months, about 60% at 12 months, and about 40% at 24 months.

Infant measurements were similar at each visit for infants exposed to progesterone gel and those exposed to placebo. Rates for congenital abnormalities and chronic morbid conditions

were higher in progesterone gel-exposed infants (5.6% and 7.2% compared to 3.8% and 5.4%, respectively for placebo-exposed infants). However, the Applicant reports that when congenital abnormalities that could not have occurred during the treatment period are excluded, and when congenital abnormalities are excluded from chronic morbid conditions, the rates are similar across groups.

Over the three assessments (at 6, 12 and 24 months), about 36-58% of the progesterone gel-exposed children and 41-62% of placebo-exposed children were screened by the DDST II. General results were similar, with only about 4-5% of the total population in each treatment arm being considered "suspect" on screening.

5.2.6 Laboratory Findings

Clinical laboratory testing was not performed in either phase 3 study.

5.3 Summary of Safety

The clinical safety database for progesterone gel included 1,119 subjects, 568 of whom received the to-be-marketed formulation of progesterone gel 8%. No maternal deaths occurred and the rates of fetal, neonatal and infant deaths were similar in both treatment arms. In general, rates of SAEs and AEs were similar across treatment arms. The infant follow-up study provided reassuring data, suggesting that there is neither a positive nor negative long-term impact of maternal treatment with progesterone gel.

---

[1] Muglia LJ and Katz M. The enigma of spontaneous preterm birth. N Engl J Med 2010; 362: 529-35

[2] Iams JD et al. The length of the cervix and the risk of spontaneous premature delivery. N Engl J Med 1996; 334: 567-72

[3] Lahra MM, Jeffery HE. A fetal response to chorioamnionitis is associated with early survival after preterm birth. Am J Obstet Gynecol. 2004; 190: 147-51